PUCRS

ESCOLA DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA
DOUTORADO EM FILOSOFIA

FELIPE MENDES SOZZI MIGUEL

EVIDENTIALISM, DEFEASIBILITY, AND THE COGNITIVE SCIENCE OF RELIGION

Porto Alegre
2019

PÓS-GRADUAÇÃO - *STRICTO SENSU*

Pontifícia Universidade Católica
do Rio Grande do Sul

FELIPE MENDES SOZZI MIGUEL

EVIDENTIALISM, DEFEASIBILITY, AND THE COGNITIVE SCIENCE OF RELIGION

Tese de Doutorado apresentada ao Programa de Pós- Graduação em Filosofia, da Pontifícia Universidade Católica do Rio Grande do Sul, como parte dos requisitos necessários à obtenção do título de Doutor em Filosofia.

Aprovada em 10 de janeiro de 2019.

BANCA EXAMINADORA:

_____

Prof. Dr. Claudio de Almeida – PUCRS

(Orientador)

_____

Prof. Dr. Rogel Oliveira – PUCRS

_____

Prof. Dr. Matthew Frise – Santa Clara University (EUA)

_____

Prof. Dr. Chris Tucker – College of William and Mary (EUA)

LIST OF FIGURES

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I am especially grateful to my dissertation adviser, Claudio de Almeida, for his steadfast support for my academic work, and to the members of the dissertation committee, Rogel Oliveira, Chris Tucker, and Matthew Frise. Tucker's and Frise's generous feedback on some of the concerns I had at an early stage of this project greatly contributed to reorient it to its current format. I am also grateful to my host at the University of Notre Dame, Robert Audi, for commenting on an early draft of chapter eight, for giving me the opportunity to present an early draft of chapter six for his Fall 2017 seminar, and for giving me so many opportunities to discuss philosophy with him. I presented an early version of chapter eight during a week-long workshop led by Justin Barrett in Salvador, in August 2017, and I am grateful to him for discussion on some of the topics of this dissertation. I have also benefited from discussion with my colleagues at PUCRS and with several philosophers whose work I discuss in this dissertation. I am especially grateful to Richard Swinburne, who provided comments on chapter four, to Andrew Moon, for discussions and for his comments on chapter six, to Paul Moser, for discussion on his views on the implications of Cognitive Science of Religion to religious epistemology. I am also grateful to Rodrigo Borges and to Rogel Oliveira for comments on an early version of chapter five. Finally, my research for this project was made possible by a CAPES fellowship and by an Oxford-Templeton scholarship.

ABSTRACT

Several claims of philosophical significance have recently been made on the basis of findings from Cognitive Science (CS) and from its subdiscipline Cognitive Science of Religion (CSR). In this essay, two such claims are evaluated. First, it has been claimed that CS indicates that evidentialism about epistemic justification is false and that CSR, specifically, indicates that religious evidentialism is false. And, secondly, it has been claimed that CSR indicates that theistic beliefs are not justified.

The first part of this essay addresses the first claim. It begins with a presentation of the current debate on the nature of evidence and on evidentialism about epistemic justification. This is followed by an exploration of some of the main findings of CSR and of three main accounts of CSR findings that are pertinent to an evaluation of the first claim. It concludes with an evaluation of the compatibility of reformed epistemology with evidentialism and of religious evidentialism with those three CSR accounts, in addition to an exploration of the prospects for the formulation of an improved account of the *sensus divinitatis* given the three CSR accounts, and with an evaluation of more general CS objections against evidentialism. It is found that CS/CSR pose no obvious threat to evidentialism/religious evidentialism.

The second part of this essay examines the second claim. It begins with an exploration of the nature of rationality and of defeaters, of Alvin Plantinga's account of defeaters, and of recent challenges to the traditional notion of epistemic defeasibility. This is followed by a presentation and examination of the twelve objections against the rationality of theistic beliefs formulated on the basis of CSR findings. It is found that CSR poses no obvious threat to the rationality of theistic beliefs.

Keywords: Cognitive Science; Cognitive Science of Religion; Evidence; Evidentialism; Religious Evidentialism; Rationality; Defeaters; Defeasibility; Theistic Beliefs.

RESUMO

Várias afirmações filosoficamente relevantes têm sido feitas com base nos resultados das pesquisas desenvolvidas no âmbito da Ciência Cognitiva (CC) e da sua subdisciplina Ciência Cognitiva da Religião (CCR). Neste ensaio, duas dessas afirmações são avaliadas. Primeiramente, tem sido alegado que a CC indica que o evidencialismo sobre justificação epistêmica é falso e que a CCR, especificamente, indica que o evidencialismo religioso é falso. E, em segundo lugar, tem sido afirmado que a CCR indica que crenças teístas não são justificadas.

A primeira parte deste ensaio aborda a primeira afirmação. Ela começa com uma apresentação do debate atual sobre a natureza da evidência e do evidencialismo sobre justificação epistêmica. Em seguida, explora-se algumas das principais conclusões da CCR e de três abordagens principais acerca dos resultados da CCR que são pertinentes para uma avaliação da primeira afirmação. A primeira parte é concluída com uma avaliação da compatibilidade da epistemologia reformada com o evidencialismo e do evidencialismo religioso com essas três abordagens da CCR, além de uma exploração das perspectivas para a formulação de uma abordagem melhorada do *sensus divinitatis* dadas as três abordagens de CCR, e com uma avaliação de objeções mais gerais da CC contra o evidencialismo. Verifica-se que CC/CCR não representam uma ameaça óbvia ao evidencialismo/evidencialismo religioso.

A segunda parte deste ensaio examina a segunda alegação. Ela começa com uma exploração da natureza da racionalidade e dos derrotadores, com uma descrição da abordagem desenvolvida por Alvin Plantinga sobre derrotadores e dos desafios recentes à noção tradicional de derrotabilidade epistêmica. Em seguida, é realizada uma apresentação e exame de doze objeções contra a racionalidade das crenças teístas formuladas com base nas pesquisas desenvolvidas no âmbito da CCR. Verifica-se que a CCR não representa uma ameaça óbvia à racionalidade das crenças teístas.

Palavras-chave: Ciência cognitiva; Ciência Cognitiva da Religião; Evidência; Evidencialismo; Evidencialismo Religioso; Racionalidade; Derrotadores; Derrotabilidade; Crenças Teístas.

INTRODUCTION

Religious beliefs and practices have been part of virtually every known human culture in history. The ubiquity of religion has led to a number of attempts to provide naturalistic explanations for the origins of religious ideas and beliefs. Anthropologist and cognitive scientist of religion Pascal Boyer (2001, chapter 1) has identified four of the main explanations that have been advanced to account for the origins and pervasiveness of religion: religion provides explanations, religion provides comfort, religion provides social order, and religion is a cognitive illusion. Proponents of the first category usually identify the origins of religious ideas and beliefs with the human propensity to seek explanations for puzzling natural phenomena and experiences (such as dreams), for the origins of things, for why there is evil and suffering, and so on. Proponents of the second category take religion to be so prevalent and pervasive because they reduce anxiety about death and about the challenges and dangers of life. The third common explanation for the origins of religious ideas and beliefs focuses on social considerations, such as their benefits in terms of holding society together and reproducing the social order, as well as in terms of their contribution to the affirmation of a moral order that binds the member of the social group. Finally, what makes religion so ubiquitous is, according to the cognitive illusion approach to religion, the human propensity to hold superstitious beliefs which are difficult to refute or even irrefutable.

Boyer believes that none of these proposals succeed in explaining the origins of religious ideas and beliefs, and he himself is a spokesman of an alternative approach to explain why religion has been such an integral component of the human experience. This alternative approach emerged as a by-product of the rise of cognitive science, the science of the human mind whose roots can be traced back to the 1960s, when linguists, anthropologists, computer scientists, and evolutionary biologists, among others, started gathering occasionally in interdisciplinary meetings to discuss the operations of the human mind. Cognitive science (henceforth CS) progressively became an empirical science that sought answers to questions about human perception, attention, memory, conceptualization,

reasoning, learning, to name a few.[1] In the early 1980s,[2] these findings began to be applied specifically to the question of the origins of religious thought, and, today, Cognitive Science of Religion (henceforth CSR) is a burgeoning field of study. Along with Boyer, other major contributors to the field of CSR are Scott Atran, Jesse Bering, Justin Barrett, Paul Bloom, Deborah Kelemen, and Robert McCauley, to name just a few.

Although CSR is still a nascent discipline and, consequently, much work remains to be done, CSR researchers have uncovered a series of fascinating findings about how our naturally developing cognition disposes us to embrace religious ideas and practices. Barrett (2012b) has identified thirteen such ideas or beliefs to which human beings are developmentally disposed. Here is Barrett's summary of the main findings that have been uncovered and remain under exploration by cognitive scientists working on religious cognition that have contributed to the ongoing effort to paint an accurate picture of the origins and permanence of religious thought and practice in human history (with references to the respective literature):

> (A) Elements of the natural world such as rocks, trees, mountains, and animals are purposefully and intentionally designed by someone(s), who must therefore have superhuman power (Kelemen 2004).
>
> (B) Things happen in the world that unseen agents cause. These agents are not human or animal (Guthrie 1993).
>
> (C) Humans have internal components (such as a mind, soul, and/or spirit) that are distinguishable from the body (Bloom 2004, 2007, 2009).
>
> (D) Moral norms are unchangeable – even by gods (Hauser 2006; Katz 2000).
>
> (E) Immoral behavior leads to misfortune; moral behavior to fortune (Jose 1990; Hafer and Begue 2005).

[1] There are several excellent introductory resources to cognitive science. For one that provides an overview of both cognitive science and cognitive science of religion, see Barrett (2011).

[2] Stewart Gunthrie's article A Cognitive Theory of Religion (1980) is taken by many to be the scholarly work that launched the field of cognitive science of religion.

(F) Ritualized behaviors such as marking off special spaces or ritual cleansings can protect from unseen hazards (including those caused by gods) (Liénard and Boyer 2006; Boyer and Liénard 2006).

(G) Some component(s) of humans that has agency (such as souls or minds) may continue to exist without earthly bodies after death (thereby becoming gods) (Cohen and Barrett forthcoming; Bloom 2004).

(H) Gods exist with thoughts, wants, perspectives, and free will to act (Guthrie 1993; Barrett 2012).

(I) Gods may be invisible and immortal, but they are not outside of space and time (Barrett and Keil 1996; Barrett 1999).

(J) Gods can and do interact with the natural world and people, perhaps especially those that are ancestors of the living, and hence, have an interest in the living. This interaction with the world accounts for perceived agency and purpose in the world that cannot be accounted for by human or animal activity (Barrett 2008; Bering 2006, 2002; Boyer 2001).

(K) Gods generally know things that humans do not (they can be super-knowing or superperceiving or both), perhaps particularly things that are important for human relations (Boyer 2001; Barrett and Richert 2003).

(L) Gods, because of their access to relevant information and special powers, may be responsible for instances of fortune or misfortune; they can reward or punish human actions (Bering and Johnson 2005; Johnson 2005; Boyer 2001; Bering and Parker 2006).

(M) Because of their superhuman power, when gods act, they act permanently, and so when they act in religious rituals, the religious ritual need not be repeated as in baptisms or ordinations (McCauley and Lawson 2002).[3]

In this essay, we will explore certain claims made by cognitive scientists and, in particular, cognitive scientists of religion of philosophical significance. More precisely, these claims are of epistemological significance. Epistemology is the sub-area of philosophy concerned with the nature of knowledge and of rational or justified belief and the conditions under which

[3] Barrett, Justin (2012b, pp. 322-23)

these properties are obtained. Among the claims of philosophical significance that have been made on the basis of findings from CS and CSR are: (1) that CS/CSR findings indicate that a particular view on the nature of epistemic justification known as evidentialism (and, hence, its correlate religious evidentialism) is false, and (2) that these findings indicate that certain religious beliefs, such as theistic beliefs, are unjustified or irrational. But how good are the objections that have been presented in the literature against evidentialism and against the rationality of theistic beliefs on the basis of findings from CS and, more specifically, CSR? This is the question that we will be attempting to answer in this essay. In the remainder of this introduction, we will lay out these two claims in more detail, and, then, we will summarize our plan for the essay.

THE FIRST CLAIM: *CS INDICATES THAT EVIDENTIALISM ABOUT EPISTEMIC JUSTIFICATION IS FALSE AND CSR, SPECIFICALLY, INDICATES THAT RELIGIOUS EVIDENTIALISM IS FALSE*

Evidentialism is, roughly, the thesis that a belief is epistemically justified only if it fits the evidence one has available when forming the belief. Recently, a number of objections to evidentialism have been presented on the basis of findings from CS.

John Greco (2006), for instance, claims that recent empirical studies show that paradigm cases of knowledge, such as perceptual knowledge, memory knowledge, and inductive knowledge, cannot be understood entirely in terms of person-level representational states, something he understands evidentialists to be committed to.

Others, such as Robert McCauley (2011), have claimed, also on the basis of paradigm cases of knowledge, that our *natural cognition* (which delivers intuitive, non-inferential, non-reflective, and instantaneous knowledge) provides us with knowledge of our environment in circumstances that go beyond what the evidence available to us at the moment indicates.

Still others, such as Justin Barrett (2004 and 2009), have objected to evidentialism on the basis of the overall picture of our belief-formation process. On this picture, evidence is always filtered and distorted by the operation of our mental tools. We never have direct access to evidence. Rather, the information used by our cognitive tools to form beliefs consists of processed information stored in our memory. This information bears little resemblance to whatever evidence we might have initially absorbed from the external

world. Thus, since we never have direct access to pure evidence when forming beliefs, our knowledge of the world cannot be said to depend on our proper response to evidence.

Since these authors formulate their objections to evidentialism on the basis of cases of knowledge without evidence, it is tempting to interpret their claims, if they are successful, as delivering only a partial defeat of evidentialism. For evidentialism, as formulated by its main proponents today (such as Earl Conee and Richard Feldman) is primarily as a thesis about justification, and, only secondarily, about knowledge. Still, most evidentialists are likely to embrace the view that knowledge requires justification and, hence, evidence. If evidence is not necessary for knowledge, then evidentialism, while it may still be the correct theory of justification, will certainly lose much of its attractiveness since knowledge is, for many, the most valued epistemic goal. Consequently, evidentialists will want to reject any view that entails that knowledge does not require evidence.

However, I believe the claims made by Greco, McCauley, and Barrett are more plausibly construed as also including the incapacity of evidentialism to account for justified belief. By invoking general considerations about the process through which our mind responds to external inputs and also about how it absorbs, processes, stores, and retrieves information, they are pointing to a deeper problem for evidentialism: it doesn't seem capable of properly accounting for how our minds are related to the world and how it responds to external stimuli. Thus, their claims seem to be more general and more damaging to evidentialism than just the question of the possibility of knowledge without evidence would suggest. If our minds don't work the way evidentialism supposes that it works, then it is not only the evidentialist view about knowledge that is at risk here, but also evidentialist accounts of epistemic justification.

In addition to these general objections to evidentialism from CS, Clark and Barrett (2010, 2011) have objected to religious evidentialism on the basis of the findings of CSR. Clark and Barrett (2010) have argued have CSR has remarkably converged with the thesis about the possibility of non-inferential justification of theistic beliefs and theistic knowledge known as reformed epistemology. Clark and Barrett (2011) construe reformed epistemology as a further development of Thomas Reid's common sense epistemology, which they take to be

"antievidentialist" (641).[4] On their reading of Reid, "we have a tendency or disposition" (646) to form certain beliefs that are justified even though we lack supporting evidence (such as beliefs about the external world and that we have other minds). The cognitive faculties involved in the formation of those beliefs "produce their effects immediately, without the evidential support of other beliefs" (646). In addition, some proponents of reformed epistemology, such as Alvin Plantinga, have argued that, if theism is true, we are probably endowed with a cognitive faculty that produces non-inferential justified theistic beliefs (and knowledge) when we find ourselves in certain circumstances. Clark and Barrett (2010, 2011) argue that CSR findings, and the thesis of the naturalness of religious beliefs that have

[4] It is not clear that Reid was in fact an antievidentialist. Here are two passages from Reid that seems to suggest that this interpretation of his views is mistaken:

"To believe without evidence is a weakness which every man is concerned to avoid, and which every man wishes to avoid" (1788, 2.20).

And

"all good evidence is commonly called reasonable evidence, and very justly, because it ought to govern our belief as reasonable creatures" (1788, 2.20).

In addition, when discussing the rationality of non-inferential beliefs, Reid writes that:

"Their *evidence* is not demonstrative, but intuitive. They require not proof, but to be placed in the proper point of view" (1785, 1.2, 42, my emphasis).

Interpretations of Reid as an antievidentialist seem to arise from his response to skepticism and to his principle of testimony, which roughly states that one is rational in her beliefs until she encounters defeaters for them. The combination of the quotations above with Reid's response to skepticism and his principle of testimony seem to suggest that Reid's epistemological views are closer to certain varieties of internalist evidentialism than to externalist views (See van Woudenberg, 2013, for discussion of Reid's views in light of the debate internalism *vs.* externalism). The eminent expert in the philosophy of Reid, James Van Cleave, believes (in personal communication) that Reid's views are closer to what is now known as phenomenal conservatism (due to Michael Huemer's work) than to anything else we have today in the epistemological scene. Michael Bergmann (2008), however, notes that Reid seems to deny the necessity of evidential fit, which would drive his views closer to proper functionalism, though not necessarily make him an antievidentialist. In fact, given what Reid had to say about evidence and its relation to rational belief, his views on the epistemology of religion may well be closer to those of Richard Swinburne, a proponent of the principle of testimony (which he calls the principle of credulity) and of epistemic conservatism, than to those of major externalist religious epistemologists, such as Alvin Plantinga. For Plantinga, while frequently saying things that put him very close to evidentialism (see Dougherty and Tweedt, 2015), has repeatedly formulated his views about warrant using expressions like "without the need of argument or evidence," though, in fairness, his main objection is to the idea that one's religious beliefs must be inferential in order for them to be rational. Thus, Plantinga's opposition seems to be to the idea that religious beliefs must be based on "inferential evidence," rather than on "evidence" *simpliciter*, in order to be rational – though the persistence of his use of "without evidence" may suggest otherwise. We will return to some of these questions in chapter four.

emerged from those findings, strongly suggest that we are in fact endowed with something resembling a God-faculty or perhaps faculties and that this remarkable convergence between these findings and Reidian (*i.e.*, antievidentialist) reformed epistemology may provide epistemic support for the latter (2010, 189).[5,6]

The first claim, then, is that the findings of CS and of CSR indicate that evidentialism in general and religious evidentialism in particular are false.[7]

THE SECOND CLAIM: *CSR INDICATES THAT THEISTIC BELIEFS ARE NOT JUSTIFIED*

In chapter three we will explore some of the main experimental findings and theories of CSR. These findings have led many to attempt to draw philosophical implications from those findings. One of the questions that has emerged is whether CSR can tell us anything about the rationality of theistic and atheistic beliefs. There are those, such as Tim Mawson, who believe that CSR doesn't have much to say about the rationality of either theistic or atheistic beliefs, for its findings can, as he put it, support both the conclusion that "God has made us for Himself, so our brains are restless unless they find their rest in Him" and the view that "our brains have 'made God' for themselves" (2014, 149). Others, however, such as Pascal Boyer (2001), Richard Dawkins (2006), and Daniel Dennett (2008), have taken the findings about the evolutionary origins of religious beliefs to support the conclusion that these beliefs are unjustified or unwarranted or even false. Others, such as Michael Murray (2009), Justin Barrett (2007), Kelly James Clark (2011), Dani Rabinowitz (2011), Joshua Thurow (2013), David Leech (2011a, 2011b), and Aku Visala (2011a, 2011b), have defended the view that the findings of CSR pose no threat to the rationality of theistic beliefs. Kelly James Clark and Justin Barrett (2010), as mentioned previously, have defended the idea that the model

[5] More recently, Matthew Braddock (2018) has argued that CSR findings provide epistemic support for theism. We will explore his claims in more detail in the final chapter of this essay.

[6] As we will see in chapter four, reformed epistemology need not be antievidentialist. In fact, two promising evidentialist formulations of reformed epistemology have been advanced recently and we will explore how these formulations can be improved in light of CSR findings.

[7] While Clark and Barrett only claim that CSR provides epistemic support for antievidentialist reformed epistemology, the earlier claims by Greco, McCauley, and Barrett seem to target the truth of evidentialism *simplicter*. And since the falsehood of evidentialism *simplicter* seems to entail the falsehood of religious evidentialism, it seems that the most accurate way to formulate the claim is in terms of CS and CSR findings indicating that both evidentialism *simplicter* and religious evidentialism are false.

of religious belief formation sketched by cognitive scientists of religion is actually supportive of the model of production of knowledge of God defended by proponents of reformed epistemology.[8] And more recently, objections to the rationality of atheistic beliefs have been proposed by Clark (2013, 2014), Barrett (2013), and Ian Church (2013). And Matthew Braddock (2018) has argued that CSR findings provide epistemic support for theism.

Thus, the second claim that we propose to explore in this essay is whether CSR can tell us anything about the rationality of religious beliefs. Since religious beliefs are varied and multifaceted, we will focus on the sort of religious belief that has attracted most scholarly debate in light of CSR findings: theistic beliefs. As a result, in the final chapter of this essay, we will present and evaluate the main arguments that have been advanced in the literature against the rationality of theistic beliefs in light of CSR findings.

WHAT LIES AHEAD

The essay is divided into two parts, corresponding to the two claims outlined above. The first part has four chapters. The first chapter presents the current debate on the nature of evidence. Five main positions on what, ontologically speaking, evidence is are sketched, in addition to the four main views on the roles evidence is expected to play. The five views on the nature of evidence are (a) that evidence consists of non-factive mental or psychological states, (b) that evidence consists of propositions, (c) that evidence is better characterized as facts, (d) that evidence is best seen as factive psychological states that are either knowledge or justified beliefs, and (e) that the correct account of the nature of evidence should be pluralistic, combining elements of all three monist positions. With respect to the roles evidence is expected to play, evidence is said to be (i) what justifies doxastic states, (ii) what rational people seek, (iii) an indicator that something is true, and (iv) a neutral arbiter between theories.

This overview of the literature on the nature of evidence is important as claim one clearly depends on how we conceive of evidence and, in fact, the objections to evidentialism from

[8] As Clark and Barrett noted, "reformed epistemology and cognitive science have remarkably converged on belief in God" (2000, 174). More specifically, the claim is that they have converged on the conclusion that belief in God is the natural state of belief for humans. Unbelief, on the other hand, requires the suppression of our natural disposition to believe in the existence of supernatural beings.

CS and to religious evidentialism from CSR that we will examine presuppose specific views on the nature of evidence. Our goal in that chapter is not to settle the debate on which view about the nature of the evidence is the correct one, but to show the possibilities an evidentialist has in formulating an account of what evidence is, something required for a complete account of evidentialism. But since most evidentialists today favor a mentalist or psychologist view of evidence, one of the main goals of the chapter is to show the variety of options this sort of evidentialist has when formulating an account of what evidence is or can be: beliefs, inclinations to believe, seemings, non-factive or non-propositional experiences, dispositions, and so forth. In the subsequent chapter, our understanding of evidentialism will presuppose this predominant view that evidence consists of mental states. And as we will see, this constitutes a crucial step in showing how an evidentialist account of the findings of CSR that is friendly to the view that religious beliefs (and theistic beliefs in particular) are not rendered unjustified by the findings of CSR can be formulated.

The second chapter presents the evidentialist thesis about epistemic justification. It lays out the schema that has to be filled by any viable evidentialism, shows the main items of such a theory besides the nature of evidence that need to be developed (evidence possession and evidential fit), and the main objections that have been advanced against evidentialism and how evidentialists have responded to them. The goal of this chapter is to show that evidentialism can be developed into a coherent and plausible theory of justification. An important aspect of the chapter is to show that, unlike the widespread assumption that evidentialism makes justification too difficult, overintellectualizing our epistemic lives, the most prominent contemporary renderings of evidentialism are fully consistent with the justification of our commonsense beliefs, a desideratum that arguably any plausible theory of justification must fulfill.

The goal of the third chapter is twofold: to present some of the main experimental findings of CSR research in support of the naturalness thesis and to present three of the main theories developed by cognitive scientists about the origins of religious beliefs: the attribution account, dispositionalism, and the preparedness theory. I lay out the main empirical findings and theories underpinning these accounts, laying the groundwork for their examination in chapter four in light of evidentialism. The criterion for selection of these

three specific theories was that these are the theories that Clark and Barrett examine in their article *Reformed Epistemology and the Cognitive Science of Religion* (2010), in which they defend the convergence of CSR and reformed epistemology. In that article they explore to what extent the god-faculty (or faculties) uncovered by CSR and the models of the *sensus divinitatis* developed by Calvin and Plantinga converge. It seems appropriate that the theories examined in this first part of the essay are, or at least include, the three theories Clark and Barrett (2010) examine. After all, the main purpose of the first part of this essay is to respond to objections to evidentialism (and religious evidentialism) from CS (CSR), and Clark and Barrett (2011) have been two of the main proponents of the objection to religious evidentialism on the basis of CSR findings. But since their paper is almost ten years old, I formulate the three theories in a way that takes more recent empirical findings and the more recent discussion of those findings into account.

In the fourth chapter (and the final one of part one), the findings of the three previous chapters are brought together with the aim of evaluating the claim that CS has indicated that evidentialism about epistemic justification is false and that, in particular, religious evidentialism should be rejected. I begin by discussing the main models of the *sensus divinitatis* and of reformed epistemology that have been proposed so far (Calvin's, Plantinga's, 2000, Tucker, 2011, and McCallister and Dougherty's, 2018). Tucker's (2011) and McCallister and Dougherty's (2018) models are formulated explicitly as evidentialist alternatives to Plantinga's proper functionalist model, and, if successful, indicate that the *sensus divinitatis* can be understood in evidentialist terms. I show, however, that if we are – as we should – to take the findings of CSR as important considerations when formulating a model of the *sensus divinitatis*, more work needs to be done. More precisely, it needs to be explored the degree to which the three theories discussed by Clark and Barrett (2010) are compatible with these evidentialist understandings of the *sensus divinitatis*. Thus, after presenting those four models of the *sensus divinitatis*, I discuss some of the most plausible formulations of the evidentialist thesis and show how some of them can deliver the correct results with respect to the three CSR accounts. With the question of whether the three CSR accounts can be given evidentialist interpretations answered affirmatively, we move to our evaluation of the objections to evidentialism advanced by Greco, McCauley, and Barrett.

Their objections are reconstructed in the form of valid arguments and all three are shown to have problematic premises. It is therefore claimed that they fail to show that CS indicates that evidentialism is false. Since the three CSR accounts are compatible with religious evidentialism and since religious evidentialism is entailed by evidentialism *simpliciter*, claim one is therefore shown to be false.

The second part has five chapters. The fifth chapter of this essay consists in an overview of the current literature on the nature of defeaters and of epistemic rationality. It begins with a brief discussion of the traditional view of rationality (which is taken to be interchangeable with the notion of justification), followed by a brief discussion of Robert Audi's (2004, 2011) alternative approach to rationality, justification, and reasonableness, and of the distinction between internal and external rationality defended by Plantinga (2000) and Michael Bergmann (2009). We then move to our discussion about defeaters. The main views on the definition and classification of defeaters and on how they are expected to function are briefly discussed. We conclude this brief chapter with a discussion of the concept of total evidence. The goal of this chapter is to provide a general understanding of what is and what isn't meant by defeaters and rationality, two concepts that will play pivotal roles in the remainder of this essay.

The sixth chapter deepens the discussion of the previous chapter with respect to defeaters. This is done by examining in more detail the account of defeaters developed by Alvin Plantinga, arguably the most detailed, nuanced, and terminologically rich account of defeaters developed so far. Among the types of Plantingian defeaters that we discuss in the sixth chapter are: warrant defeaters, proper-function-rationality defeaters, Humean defeaters, purely alethic rationality defeaters, potential defeaters, defeater-deflectors, defeater-defeaters, neutralizing defeater-defeaters, intrinsic defeater-defeaters, extrinsic defeater-defeaters, partial defeaters, and "optimistic overriders." No doubt an impressive panoply of defeaters!

Plantinga's account of defeaters is particularly relevant for the purposes of this essay in that it was developed, to a large extent, in response to criticisms of his Evolutionary Argument against Naturalism (EAAN). The EAAN is an argument to the effect that those who believe the conjunction of two propositions, namely, that naturalism and evolution are true, have a

defeater for the reliability of their cognitive faculties. Since most, and the strongest, CSR objections to the rationality of theistic beliefs come in the form of claims that the cognitive mechanisms that produce theistic beliefs are not reliable, it is particularly instructive, for the purposes of the second part of this essay, for us to examine Plantinga's responses to objections to his EAAN. Along the way, we explore other questions related to Plantinga's views on epistemic defeasibility (*e.g.*, whether he has turned his theory of knowledge into a defeasibility theory) and four objections to his views.

The idea that justification and knowledge are defeasible has become one of the cornerstones of contemporary epistemology. As a result, most theories of justification or knowledge have sought to make room for the notion of defeat. Recently, however, several challenges to this notion have emerged. In the seventh chapter we explore some of the alleged problems with the notion of defeat in epistemology as developed by Maria Lasonen-Aarnio and Max Baker-Hitch and Mathew Benton. We conclude by suggesting that two of the evidentialist theories of epistemic support that we examined in the first part of this essay can potentially escape those problems, which increases the plausibility of these accounts of evidential support in comparison with the externalist and internalist alternatives that face the problems raised by Lasonen-Aarnio and Baker-Hitch and Benton. Our primary goal in discussing these criticisms of the traditional notion of defeasibility and in suggesting that there are evidentialist theories that can plausibly escape those criticisms is to provide a general picture of the current debate on the nature and varieties of epistemic defeat, complementing the discussions about epistemic defeat of the previous chapters.

While the first part of the essay is primarily concerned with justification or rationality in light of one's total evidence, the second part is concerned more specifically with whether there are CSR findings that could function as genuine (rather than misleading) defeaters for the justification or rationality of one's theistic beliefs were those findings and interpretations of them to become part of one's total evidence. On evidentialism, one has propositional justification to believe *p* when, roughly, her total evidence supports *p*. But perhaps there are (genuine or misleading) counterevidence *e* outside of her total evidence (*i.e.*, outside of her mental life, in the case of mentalist evidentialism) that would prevent her from believing *p* rationally were *e* to become part of her total evidence. While we may never know whether a

given piece of evidence *e* is truly genuine (there might always be defeaters that we aren't aware of), I am using the term genuine to indicate that a given piece of evidence *e* cannot be undermined or shown to be false by current philosophical and scientific research. With this distinction in mind, we evaluate, in the ninth and final chapter of the essay, twelve objections to the rationality of theistic beliefs that are found in the literature. These objections could potentially constitute defeaters for the rationality of theistic belief. If there are successful CSR objections to the rationality of theistic belief, then our philosophical and scientific evidence *suggest* that there are genuine defeaters for those beliefs.

The objections under evaluation vary in structure and strength. While some are patently unsound arguments, others have a valid structure and premises that are not obviously false. It is shown that even the strongest objections face serious problems that strongly suggest that they do not constitute (genuine) defeaters for the rationality of theistic beliefs. The twelve objections are: The Natural Explanation Objection I, the Natural Explanation Objection II, the Neural Substrate Objection, the By-product Objection, the Religious Utility Objection, the Inherited Beliefs Objection, the Lack of Proper Causal Relationship Objection, the False Positives Objection, the Mutually Exclusive Beliefs Objection, the Simplicity Objection, the Problem of Natural Non-Belief, and the Confirmation Bias Objection. Thus, the main conclusion of this essay is that the two claims under consideration – that religious evidentialism is rendered false by cognitive science of religion and that theistic beliefs are (genuinely) defeated by the CSR findings and theories – fail to accomplish their intended aims. If these results are correct, we can say that, as far as CSR is concerned, religious evidentialism and theistic beliefs remain undefeated.[9]

[9] As we will see in chapter nine, the sense in which theistic beliefs remain undefeated is in terms of full defeat and in terms of the scientific and philosophical research *suggesting* that there are genuine defeaters for the justification of theistic beliefs from CSR. For findings of CSR can defeat justification by entering one's total evidence as misleading evidence, and they can provide (genuine) grounds or evidence (in certain limited circumstances to be examined in the final chapter) for some reduction of justification, thus providing partial defeat for the justification of theistic beliefs.

PART I

EVIDENCE

Evidence is a fundamental and pervasive notion in human life. It is present both in complex intellectual activities and in the simplest and most trivial ones. Both the astrophysicist looking for an answer for the origin of black holes, and the detective trying to solve a crime, seek, evaluate, and respond to evidence. The same can be said of the historian in search for the correct explanation for the collapse of a particular civilization, or of the philosopher trying to identify which position in a particular philosophical debate that has been going on for millennia (*e.g.*, on free will, or on the nature of the human person, or on the definition of knowledge) is the correct one. Likewise, a consumer seeks evidence of the quality of tomatoes in the grocery store by observing their color, their smell, the sensation of their texture and softness.[10]

In this chapter, we will present the contemporary debate on the nature or ontology of evidence and on the roles that the notion of evidence is expected to play in the relevant human activities. Among the issues that will be discussed are: Does evidence consist of propositions, mental states, facts, or a combination of these options? Can there be false evidence? All evidence consists of knowledge? Evidence is what justifies beliefs? Or is it what rational people seek? Or is it perhaps an indicator that something is true? Or, still, a neutral arbiter between theories? Or, perhaps, a combination of these four roles?

With respect to the nature of evidence, that is, what it consists of, ontologically speaking, six positions can be found in the recent literature.[11] For some, evidence consists of propositions (Mcgrew, 2011, Dougherty, 2011, Neta, 2008). For others, evidence would be better characterized as facts (Dancy, 2002, Kelly, 2008, Neta, 2018). Others argue that evidence consists of psychological states (Conee and Feldman, 2008, McCain, 2014, Turri, 2009). There are those who take evidence to be factive psychological states, disagreeing among themselves on whether these states consist of knowledge (Williamson, 2000) or justified beliefs (Littlejohn, 2012). Others prefer to adopt a pluralistic approach to the nature of evidence, arguing that the concept of evidence must capture elements of the three monistic

[10] Example found in McGrew (2011).

[11] I follow here Mitova's classification (2015).

positions, *i.e.*, propositionalism, psychologism, and factualism (Kelly, 2008, Rysiew, 2011). A position that claims to be able to bring all these elements together into a unified account of the nature of the evidence is the one advanced by Mitova in terms of what she calls “truthy psychologism” (Mitova, 2015). Let us take a look at each of these positions in more detail.

Factualism

The notion of evidence generally used in non-philosophical contexts, such as in scientific contexts and in everyday settings, is in terms of objects or facts. Evidence, in this conception, is publicly available and could be observed and studied by researchers (Swinburne, 2011); it is something that could be pointed out, or even, as Kelly (2006) illustrates, put in a plastic bag. Examples of evidence in this regard would be documents and objects studied by historians, artifacts excavated by archaeologists, laboratory observations by the chemist, the physicist, or the biologist, the knife with blood found at the crime scene and collected for examination by the detective in charge of the case, etc.

Some philosophers (Williamson, 2000; Lyons, 2015) consider that it is not the objects that could be stored in a plastic bag that would by itself play the role of evidence in these examples. The knife stained with blood itself is not evidence, but *the fact that it was found among the suspect's belongings*. At any rate, evidence in terms of facts or objects would consist of something objective and independent of the human mind that are, at the same time, distinct from true propositions. Evidence as facts consist of, as Lyons puts it, "states of the world that make the relevant propositions true" (2015, 2). Conee and Feldman (2008) contrast the notion of evidence that they favor, which they call justifying evidence, with this notion of evidence as something publicly available, independent of the human mind, and indicating the truth of something – what they call scientific evidence. The smoke coming out of the chimney and indicating fire in the hearth, or red spots indicating measles, are examples of evidence in this sense – they cannot be placed in a plastic bag –, but they are not psychological states, or an abstract object like a proposition.

Jonathan Dancy argues that evidence consists of reasons and that reasons, both motivating and normative, are facts. "We normally," argues Dancy," try to explain an action by showing that it was done for good reason [. . .] But psychological states of the agent are the wrong

sorts of thing to be good reasons." Consequently, he adds, "A believing cannot be a good reason for acting, because a good reason for acting is a reason that favours acting, and such things [. . .] are states of the world, not psychological states of the agent" (2000, 106). Therefore, evidence consists, according to Dancy, of facts, not mental states or even propositions. Mental states cannot be considered as evidence since this would lead to an ontological separation between motivational (psychological) and normative (factive) reasons, which Dancy rejects – and, furthermore, by acting for reasons we are, he claims, motivated by facts and not by mental states about facts. And propositions cannot be evidence, for they are entities that, in Dancy's view, "are too thin or insubstantial to be able to make an action wrong" (115).

Neta (2018) also advocates an approach to the nature of evidence that relies on the notion of facts. He endorses Goldman's (2009) position on evidence (which we will examine below), according to which a proposition forms part of the evidence set of a given subject *S* at time *t* if the agent has non-inferential propositional justification for this proposition. Neta, however, adds a factivity condition to it: our evidence consists of all and only facts which are non-inferentially propositionally justified for us. Facts, according to Neta, are "what is manifest to you." And "A fact *p* is manifest to an agent A if the very fact that *p* constitutes A's propositional justification for *p*" (15). For example, our propositional justification for "I exist" or "I am conscious" would consist in the very fact that we exist and are conscious. Hence, our evidence consists of facts that we cannot reject. And the facts that we are non-inferentially justified in believing are "simply the facts that [we are] in a position to know non-inferentially" (16).

Psychologism

On the psychologist's view advocated by Moser (1989), Conee and Feldman (2008), McCain (2014), and Turri (2009), among others, evidence consists of nonfactive mental states. Among the types of mental states that proponents of psychologism point to as evidence are

introspective, perceptual, and memorable experiences, and, in a somewhat more controversial way, intuitions, rational insights, beliefs, and inclinations to believe.[12]

The psychologist position is usually motivated by examples such as the following, presented by Conee and Feldman (2008, 87):

> Your evidence for the proposition that is warm where you are typically includes your sense of warmth, your evidence for the proposition that you are frustrated by being stuck in the heat in a traffic jam typically includes a palpable sense of your own frustration, your evidence for the proposition that the car in front of yours in the traffic jam is red typically includes your visual experience of how the cars looks, and so forth.

Although many psychologists accept beliefs as constituting evidence, Conee and Feldman (2008, 87), perhaps the two main proponents of the psychologist thesis today (which they call "mentalism"), argue that only experiences are ultimate evidence, and that beliefs are evidence only in a derivative way.[13] For Paul Moser, evidence is an indicator of truth, an indicator that a particular proposition is true. This indicator could be propositional or non-propositional (Moser, 2008, 35-6). Ultimate evidence, however, consists, according to Moser, in non-conceptual and non-propositional mental states (1989, 106) that justify

[12] See Climenhaga (2017) for the defense of the thesis that philosophers rely on intuitions as evidence and discussion of the literature on whether intuitions are evidence. For some philosophers, the justification of *a priori* beliefs depends on a rational insight into the truth of propositions expressing necessary truths, and perhaps these insights might, like intuitions, be considered as evidence. See Bonjour (1998) for discussion and defense of rational insight as a source of justification. See Lycan (1988) and Swinburne (2001) for discussion of the idea that beliefs are evidence, and Sosa (2007) and Swinburne (2001, 2011) for discussion of the idea that inclinations to believe are evidence. We will return to this in chapter four.

[13] However, Conee and Feldman imply that the evidential role played by experiences does not preclude evidence from being propositional: "We have argued that evidence includes experiences. This seems most intuitive to us, and it appears to have no significant cost. But we note that the general doctrines of evidentialism do not depend on any evidence being non-propositional "(2011, 321). And above all, "We need not deny that all evidence is propositional. A visual experience as of something blue against a white background might consist in awareness of propositions to the effect that certain visual qualities are arranged in a certain configurations "(2008, 101). However, as Dougherty puts it, "it is not enough that evidence be proposition*al*. The state of awareness mentioned might have a proposition as a content, but the state of awareness is not in itself a proposition, and therefore cannot stand in the relations necessary to play the key functional role of evidence. The content itself certainly can, and that is the thesis I'm defending "(2011, 229).

beliefs in certain propositions when they are part of the best explanation for the existence of those states or psychological contents (1989, 91 -99).[14]

However, some philosophers see more vices than virtues in psychologism. For Dancy (2011), we act for reasons when we are being motivated by facts that underlie these reasons, not by our mental states that represent these facts. Since an agent could not be motivated to act by a psychological state, but only by facts, *i.e.*, by states of the world, evidence then cannot consist of mental states. Kelly (2008), on the other hand, expresses doubts about the capacity of psychologism to "provide us with an epistemic foothold in the world sufficient to underwrite the knowledge we ordinarily take ourselves to have." (2008, 945). This lack of secure contact with reality would, according to Kelly, make psychologism particularly vulnerable to skeptical threats. Psychologists such as McCain (2014) and Moser (1989), however, argue that skepticism is a real problem and that it is a merit of theories that adopt the psychologist or mentalist view of evidence that they recognize it and seek to respond to it with the resources of psychologism – in the case of these two authors, through the idea that the best explanation for certain characteristics of our mental states is that the objects of the external world cause our mental states about such objects.

Kelly (2006, 2008) and Williamson (2000) argue that a conception of evidence purely in terms of mental states is problematic because it is difficult to reconcile it with the notion of evidence which seems to emerge from the scientific practice. As Williamson puts it,

> If one's evidence were restricted to the contents of one's own mind, it could not play the role that it actually does in science. The evidence for the proposition that the sun is larger than the earth is not just my present experiences or degrees of belief (2000, 193).

And Williamson briefly explores whether the inclusion of other people's evidence could remedy the problem of obtaining the evidence involved in scientific practice:

[14] This condition of justification in terms of the best explanation for the existence of certain mental states or psychological contents is just one of the several options in terms of evidential support that will be discussed in chapter two. For an alternative account purely in terms of direct acquaintance, see, for example, DePoe (2012).

> If the evidence is widened to include other people's experiences or degrees of belief, or my past one's, then my identification of it becomes even more obviously fallible. In any case, this does not seem to be the best widening; it is more plausible that the evidence for a scientific theory is the sort of thing which is made public in scientific journals. If evidence is like that, our identification of it is obviously fallible (193).

Thus, for Williamson – who, as we shall see, argues that evidence consists of known propositions – the psychologistic conception of evidence should be rejected for failing to grasp the way in which the notion of evidence is employed in scientific practice. Conee and Feldman deal with this question by recognizing that there are two notions of evidence. The first, which consists of reliable indicators of states in the world and which would be publicly accessible, is called by them "scientific evidence." The second, which consists of reasons to believe, something that serves as a justifying basis for beliefs, is called by them "justifying evidence." It is this second type of evidence that serves as the ultimate, fundamental evidence for Conee and Feldman, in the sense that "one can have scientific evidence without having any reason at all to believe what the scientific evidence supports" (2008, 85). Thus, the separation of scientific evidence from justification, far from signifying that they would be segregated into specific compartments, means that they

> are connected in a way that makes sense of the use of "evidence" for both and gives epistemic instrumental value to scientific evidence. Gaining evidence, E, for proposition P, is often the most practical way for one who knows the association of E with P to gain justifying evidence for P. This is a main way in which scientific evidence fosters knowledge (2008, 86).

Williamson uses the term phenomenological conception of evidence to refer to psychologism or mentalism, which is the dominant tradition in philosophy since Descartes when it comes to the nature of evidence. As we have seen, according to this conception of evidence, our evidence consists of mental states, *i.e.*, experiences and, perhaps, doxastic states. Other important figures in the history of philosophy who have championed positions

in line with the phenomenal conception are John Locke, George Berkeley, David Hume, Bertrand Russell, W. V. O. Quine, and most recently David Lewis. Russell conceived evidence as consisting of sensory states. For Quine, evidence would consist of stimuli from our sensory receptors. And according to Lewis, our sets of evidence include all our "perceptual experiences and apparent memories" (1996, 424).

Propositionalism

Propositionalism says that only propositions can be considered as evidence. A fundamental motivation of the propositionalist thesis stems from the idea that the evidence for a belief must be in a logical or probabilistic relation to the believed proposition and that only propositions can play this role. Propositionalists tend to view as metaphorical discourses that treat as evidence objects such as knife or stained clothing. These objects underwrite beliefs in propositions, such as "the knife belongs to the butler" or "the blood-stained clothing belonged to the victim." Timothy Mcgrew (2011) describes the advantages of the propositionalist position vis-a-vis the factualist as follows:

> The position that evidence is, in the strict sense, always propositional, has many attractions. Propositions can be believed or disbelieved, but fingerprints cannot; to say that one disbelieves a fingerprint seems to be a shorthand for saying that one does not believe that the fact that this fingerprint is present (a proposition) indicates that the defendant is guilty (another proposition). Propositions can stand in logical relations to one another, can entail each other, can be negated, conjoined, and otherwise logically manipulated. But we can no more create a disjunction between a proposition and a fingerprint than we can divide the number seven by a banana (2011, 59).

Mental states cannot be manipulated logically and therefore cannot play the role of evidence. Williamson formulates three arguments in defense of propositionalism from the nature of explanation, the nature of probabilistic relations, and the nature of deductive relations.

According to the first argument, what is explained by hypotheses, that is, the evidence, has a propositional nature. In order to substantiate this thesis, Williamson presents some considerations, among them that explanations have as structure "X is the case because Y," in which the sentences before and after "because" express propositions. In explaining an event – World War II, for example – what we do is not to explain the event itself, but "why it occurred, or had some distinctive feature" (2000, 195).

According to the second argument, since evidence confers probability, it must have probability. But, "what has a probability is a proposition; the probability is the probability *that* "(2000, 1996). Both what confers probability and what receives probability must be propositional (as Dougherty notes, 2011, 227: "What it would be for an experience to have a probability is mysterious"). So evidence must be propositional.

The third argument is based on the idea that logical relations cannot involve experiences, but only propositions. Only propositions can imply and be implied. Inconsistency is therefore a relation between propositions. Since, at times, evidence eliminates hypotheses due to inconsistency between them, evidence is propositional.

Dougherty (2011, 227-8) formulates Williamson's three arguments in favor of propositionalism as follows:

The Explanatory Argument

1. Evidence is the kind of thing which hypotheses explain;

2. But the kind of thing which hypotheses explain is propositional;

3. Therefore, evidence is propositional.

The Probabilistic Relation Argument

1'. Evidence has probability;

2'. Only propositions have probability;

3 '. Therefore, evidence consists in propositions.

The Deductive Relation Argument

1''. Sometimes, evidence rules out hypotheses by being inconsistent with them;

2''. But inconsistency is a relation defined over propositions;

3''. Therefore, evidence consists in propositions.

Conee and Feldman respond to Williamson's first two arguments in favor of propositionalism. On their view, it is a mistake to think that only propositions can figure in explanatory and probabilistic reasoning. Experiences, whether they are "an event, a states of affairs, or some other non-propositional entity" (2008, 102) could also figure in inferences to the best explanation and in probabilistic inferences. Non-propositional experiences could, therefore, figure in probabilistic reasoning. The phrase "probability on the evidence" could, as Conee and Feldman put it, be easily understood as "the probability on the proposition that the evidence occurred" (2008, 102). In the case of inferences to the best explanation,

> a request to explain the evidence consisting in an experience can equally be understood as a request to explain why the experience occurred, obtained, or had some other feature (Ibid).

Williamson argues that in seeking to explain, for example, World War I, we would be explaining propositions as to why the war took place. But Conee and Feldman respond by saying that offering explanations for the First War does not consist in explaining propositions, but actually in explaining the "occurrence of the war that the proposition asserts to have occurred" (2011, 322). Although propositions represent explanatory reasoning, what is being explained are events, not propositions. McCain illustrates this as follows:

> When Sasha describes a sunset for Sara, she will do so by using sentences that express propositions. She will say things like 'there were few clouds,' 'the sky along the horizon had a purplish hue,' and so on. The fact that Sasha describes the sunset using propositions does not mean that she is actually describing a proposition – clearly, she is not. Sasha is describing a sunset. Similarly, the fact that an explanation is itself put in terms of

> propositions does not entail that what is explained must be a proposition (2014, 14).

McCain (2014, 13-15) responds to Williamson's first argument by emphasizing the problems posed by Conee and Feldman and adding that Williamson works with a deficient notion of the nature of explanation. McCain notes that most of the approaches to the nature of explanation take events, not propositions, as the explanandum. Williamson's third argument suffers, according to McCain, from the same problem raised by Conee and Feldman for the first and second arguments: propositions illustrate explanatory and probabilistic relations, but the explanandum consists actually of events. Likewise, the fact that logical relations can be understood in terms of propositions does not mean that evidence consists of propositions. Logical relations could be equally understood, *mutatis mutandis*, in terms of events, in terms of the occurrence of evidence.

Not all propositional approaches presuppose that evidence is knowledge (Williamson's influential proposal, which we will see below), or even factive. In fact, the notion of evidence in terms of propositions seems fully compatible with both internalist and mentalist approaches to justification. Dougherty (2011), for example, argues that evidence does not consist of propositions in themselves, as Williamson, for example, argues, but in the propositional content of experiences. In this case, the evidence is a proposition that describes the experience. In contrast, experiences are not evidence, but that which provides evidence. Evidence, then, consist, on this proposal, of propositions that are provided by experiences – because they consist of the content of the propositions.

Similarly, externalists who reject the idea that evidence is knowledge as well as the thesis of factivity of evidence could also adopt a propositionalist approach to evidence. One approach that seems to follow in this direction is that of Goldman (2009). For Goldman, a proposition is part of the evidential set of a given subject *S* at time *t* if the agent has non-inferential propositional justification for this proposition. Goldman speaks of the experience of seeing a computer as providing justification for believing the proposition "there is a computer screen before me" and that this proposition would be "an item of evidence" (2009, 86). He speaks of experience as "[providing] with evidence for many propositions." He adds that "it might

be debated exactly *which* propositions they are evidence for – physical-objects propositions, internal-states propositions, and so on. . ." (2009, 90). However,

> [and] in virtue of these perceptual experiences, he is non-inferentially justified in believing any propositions, then, according to E=NPJ [non-inferential propositional justification, to be treated in more detail below], these propositions are items of evidence for him (Ibid.).

And, as we have seen, Neta (2018) endorses Goldman's position on evidence, but adds a condition of factivity: our evidence consists of all and only facts which are non-inferentially propositionally justified for us.

Factive-State Psychologism

According to Williamson's factive-state psychologism (2000), evidence consists of known propositions that are psychological states. Williamson's most original epistemological theses are that knowledge (a) is a primitive notion, that is, non-analyzable, (b) it is a mental state (hence the thesis has become known as "knowledge first" epistemology (or K1 epistemology), and (c) that evidence consists of the propositions known by the subject (E = K). Evidence, because it is knowledge, as conceived by the K1 thesis, is something factive – and therefore propositional, albeit psychological.[15]

Williamson presents several arguments in defense of E = K. According to one of them,

A. When we prefer a hypothesis *h* to an hypothesis *h** because h explains our evidence *e* better than *h** does, we are standardly assuming e to be known; if we do not know *e*, why should *h*'s capacity to explain *e* confirm *h* for us? (2000, 200).

Another argument says that

[15] Littlejohn (2012) advocates a modified version of factive-state psychologism, in which evidence consists of justified beliefs, with justification implying truth. See Mitova (2015, 2, note 2) for the attribution to Littlejohn of state-factive psychologism.

B. It is hard to see why the probability of *h* on *e* should regulate our degree of belief in *h* unless we know *e*. Again, an incompatibility between *h* and *e* does not rule out *h* unless *e* is known (2000, 200).

Still another,

C. If I observe the truth of *e* and then I forget all about it, my evidence no longer includes *e*. It is hard to see how evidence could discriminate between hypotheses the way we want it to if it did not have to be known (2000, 201).

And finally,

D. If evidence required only justified true belief, or some other good cognitive status short of knowledge, then a critical mass of evidence could set off a kind of chain reaction. Our known evidence justifies beliefs in various true hypotheses; they would count as evidence too, so this larger evidence set would justify belief in still more true hypotheses, which would in turn count as further evidence . . .. The result would be very different from our present conception of evidence (2000, 201).

In response to Williamson, Alvin Goldman (2009, 86-89) presents an alternative position that understands the nature of the evidence in terms of non-inferential propositional justification (NPJ)[16] and seeks to show how this thesis (E = NPJ) responds as well or even better to the considerations put forward by Williamson in defense of E = K. An example of NPJ is of a person having an experience of seeing a computer screen before her. Since the person has propositional justification for believing that "there is a computer screen before me," that proposition is evidence for that person. In response to (A), Goldman says that when we are justified in believing a proposition, we usually believe that we are justified in this belief and assume that the proposition is true. And all that would be necessary to make sense of the idea that *h* explains *e* better that *h** and that *e* confirms *h* would be the belief in the truth of *e*. Likewise, the belief in the truth of *e* would be, according to Goldman, sufficient for the understanding of the phenomena presented in (B): what would regulate our degree of belief in *h* would be our degree of confidence in the truth of *e*, and belief in the truth of *e* would be

[16] By propositional justification Goldman means justification to believe, even if there is no belief. If there were beliefs, the justification in question would be doxastic.

sufficient for the incompatibility between *h* and *e* eliminate *h*. With respect to (C), Goldman points out that the idea that my set of evidence does not include *e* due to forgetfulness is easily explained by E = NPJ: forgetfulness of *e* leads to the loss of non-inferential propositional justification for belief in *e*. Finally, the problem with Williamson's solution to (D), according to Goldman, is that the threat of chain reaction is only possible in cases of inferential justification, which would affect the E = K principle itself, which is not restricted to non-inferential evidence, but would not affect NPJ, an explicitly non-inferential condition.

Comesanã and Kantin (2010) present two problems that they consider decisive against E = K: if this principle is true, there are no cases of Gettier, and the principle of closure of justification is false. Regarding the first problem, Gettier cases show that knowledge requires something beyond justified true belief. In such cases, true belief is often justified by a false proposition.[17] But if the proposition that contributes to the justification of the belief is false, it is not knowledge, and therefore could not be evidence, given E = K. The existence of Gettier cases, therefore, because they involve false evidence, indicates that there is evidence that is not knowledge. The second problem arises from the denial by the defender of E = K of closure principles of justification such as "if S justifiably believes that *p* and *p* implies that *q*, S is justified in believing that *q*".[18] An implication of E = K is that unless you know that *p*, instead of merely justifiably believing that *p*, you are not justified in believing that *q*. Thus, the defender of E = K, in the absence of convincing answers to these problems, seems to have to deny the existence of Gettier cases and the truth of the closure principle. But, as Comesanan and Kantin put it, "there are Gettier cases, and the closure principle is true. Therefore, evidence isn't knowledge" (2010, 447).

Given that knowledge is factive and evidence is knowledge, evidence is then factive, on E = K. Therefore, in defending E = K, Williamson is indirectly advocating the factivity of evidence. Still, Williamson offers two specific considerations in support of the factivity thesis. The first points to the fact that the existence of false evidence would lead to the exclusion of truths:

[17] *E.g.*, in the original Gettier case of the ten coins, one of the propositions that justify the conclusion that whoever has ten coins in his pocket got the job is the false proposition that Jones got the job.

[18] This is a rather simplified version of the principle and probably inadequate. However, it is sufficient to illustrate how principles of closure of justification in general seem to pose a serious difficulty for E = K. For a more detailed discussion of the closure principles, see Hawthorne (2004).

> [If] one's evidence included falsehoods, it would rule out some truths, by being inconsistent with them. One's evidence may make some truths improbable, but it should not exclude any outright. Although we may treat false propositions as evidence, it does not follow that they are evidence. No true proposition is inconsistent with my evidence, although I may think that it is (2000, 201).

The second consideration comes from the idea that factivity would explain the fact that we adjust our beliefs to the evidence:

> [By] adjusting our beliefs to the evidence. . . [we are] adjusting them to the truth. Although true evidence can still support false conclusions, it will tend to support truths (2000, 202).

Goldman (2009, 88) responds to Williamson's first attempt to defend facticity by conceding that there are no true propositions inconsistent with our evidence and by adding that this would not be a problem for the advocate of the existence of false evidence as this would not imply permanent exclusion of truths on the basis of two considerations: the existence of misleading evidence and the possibility that something that is evidence at one point may fail to be so at another time, as when evidence is forgotten (and Williamson acknowledges that there is misleading evidence and that evidence can be forgotten, 2000, 218-19). That is, according to Goldman, although there is false evidence, this would not imply the necessary and permanent loss of true evidence; such loss can be prevented by these two considerations. As for Williamson's second defense of the factuality of evidence, Goldman says that positions that equate evidence with justification, as is the case with NPJ, suppose that justified propositions are probably true. In this conception, therefore, belief adjustment would occur with respect to what is probably true, and this, according to Goldman, would be a result as adequate as that provided by E = K, since, as Williamson acknowledges in the quoted passage of the second defense, E = K also does not eliminates the possibility that false conclusions are supported by the evidence.[19]

[19] See Neta (2018, 12) for discussion of possible problems with Goldman's criticisms of Williamson's defense of factivity.

Gettier cases, as we have seen, pose a serious problem for E = K. This problem for E = K is emphasized by Arnold (2013), which presents it as a more general problem for the factivity of evidence. The existence of what became known as benign falsehoods would show the possibility of knowledge from false propositions and, therefore, the existence of false evidence. Benign falsehoods are false propositions that are causally and evidentially essential parts in inferences that constitute knowledge (de Almeida, 2018). For example, a teacher takes 100 copies of a handout for a lecture. When he arrives in the auditorium, he counts the students, arriving at number 53. From that he infers that he has enough handouts with him. However, he counted one extra person – there are 52 people in the auditorium. Intuitively, the teacher knows that he has enough handouts with him (de Almeida, 2018, 295). The premise of this inference, however, is false, something that contradicts the factivity thesis of evidence. Therefore, in some cases, evidence is false.

Another problem with the factivity of evidence stems from the supposed possibility that falsehoods increase the probability of propositions and thus justify beliefs. On the traditional view advocated by Fantl and McGrath (2009), an important difference between justification and knowledge is that the former, unlike the latter, does not require truth. A falsehood could then be a reason that would justify a particular belief or action. And, since what is justifiably believed must belong to our sets of evidence, falsehoods could be evidence.

Littlejohn (2012) seeks to defend the factivity of evidence from Arnold's and Fantl and McGrath's objections. According to Littlejohn, in cases of benign falsehoods "the subject acquires knowledge because treating something *as if* it is evidence is a safe way of forming beliefs" (2012, 159, emphasis in the original). The problem with Fantl and McGrath's critique, on the other hand, is that the case used by them to illustrate the thesis that falsehoods constitute evidence involves the notion of factual error, rather than normative mistakes, and that factual errors cannot be considered reasons to believe or act, on Littlejohn's views.[20]

[20] Additional arguments and counter-examples against the factivity of the evidence and responses to defenses of factivity presented by Littlejohn can be found in McCain (2014, 23-27).

Many proponents of non-factual evidence take the New Evil Demon Problem (NEDP) as a strong consideration against the thesis that there isn't false evidence. According to NEDP, *S* and his counterpart *S**, who is in a scenario of manipulation of their mental states by a malignant demon, possess "exactly the same experiences, apparent memories, and intuitions, and in both worlds [. . .] go through exactly the same processes of reasoning, and form exactly the same belief" (Wedgwood, 2000, 349). The prevailing intuition in this case is that *S* and *S** have the same evidence and are in the same epistemic situation in terms of justification. However, if evidence is factive, *S* and *S** do not have the same evidence and are not in the same epistemic situation in terms of justification, since everything that goes on in the mental life of *S** is false. T Therefore, the proponent of factivity would have to reject the rather plausible thesis that *S* and *S** have the same evidence and are in the same epistemic situation in terms of justification.[21, 22]

In addition to the problems presented above for Williamson's factivity-states psychologism, it is important to note the existence of four additional problems: (a) the difficulty in understanding how propositions, which are abstract objects, could be psychological states, (b) the existence of counter-intuitive results in Williamson's thesis about what constitutes evidence, and (c) the problem that non-doxastic sensory states and illusions pose to the idea that evidence is factive.

The first problem stems from the fact that, for Williamson, knowledge is a state of mind and evidence constitutes, at the same time, propositions and knowledge – and therefore mental states. Propositions, however, are widely held to be abstract objects and therefore objective entities, independent of the human mind and unable to enter into causal relationships. But if evidence is knowledge, then evidence consists of mental states, and therefore we have the incredible thesis that evidence consists of abstract objects located in the human mind (see McCain, 2014, 21).

---

[21] The principle E = K therefore produces the result that *S* and *S** do not have the same evidence. As *S** is unaware of the demonic scenario, he believes falsehoods. Williamson regards this as a desirable result, which would provide an adequate response to the problem of external world skepticism (2000, chapter 8).

[22] Andrew Moon (2012) argues that the NEDP would equally affect virtually all versions of internalism, and hence psychologism. Kevin McCain responds to Moon's argument in (2015) and Moon responds to McCain's objections in (2015).

The second problem, as McCain (2014, 22) also points out, is that if E = K is true, information that should be counted as evidence in the formation of justified true beliefs are not considered evidence. For example, if E = K is true, then in the famous case of the barn we have a situation where, when *S* sees a false barn and forms the belief that "there is a barn," this proposition, due to the factivity of knowledge, cannot be considered as evidence. Now imagine a counterpart of *S*, *S**, which is in the same scenario, with the difference that the barn observed by S* is not fake. In this case, we have the counter-intuitive result that, unlike what happens to *S*, the proposition "there is a barn" would be evidence for S*.

Finally, Conee and Feldman (2008, 103-4) and Goldman (2009, 89) point to the problem that E = K implies that non-doxastic sensory states would not be evidence. If evidence constitutes knowledge and knowledge implies belief, then there could be no evidence without belief. Thus, perceptual appearances could not constitute evidence, which seems quite implausible. Williamson anticipates this problem and seeks to respond to it by saying that what would constitute evidence in cases of sensory appearances would be known demonstrative propositions. Williamson acknowledges that known propositions would not be able to exhaust the experiential richness of our vision of, for example, a pointed mountain. He suggests, then, that the experiential evidence in this case would come in the form of the person indicating with the finger the shape of the mountain: "It is this form." (2000, 197-8). Conee and Feldman raise doubts about the plausibility of Williamson's response, among them the existence of the additional problem of illusory perception (2008, 103-4). Propositions about illusions, such as, for example, about a mirage in a desert, would be false and therefore would not be evidence given the factivity of evidence derived from the formula E = K.

After contrasting E = K with NPJ in terms of the various arguments provided by Williamson in defense of E = K, Goldman concludes that:

> Williamson does not offer a compelling rationale for the E = K thesis. A preferable view, notably different from E = K, has been outlined, which accounts for all the intuitive “data” Williamson adduces for E = K. No doubt, there are many other possible views as well, intermediate between E = K and NPJ. The main point,

however, is that no convincing arguments has been given for E = K. (2009, 90).[23]

Pluralism and Truthy Psychologism

Pluralists, such as Kelly (2006) and Rysiew (2011), argue that there is no single notion of evidence capable of accounting for the various roles played by the notion of evidence. Thus, instead of arguing for a single notion of evidence, they argue that the best thing to do is to recognize the existence of at least two notions of evidence (one fundamentally psychological and one fundamentally factive). Kelly (2006) presents four roles played by the concept of evidence and notes that it is not possible that the concept present in the first two is the same present in the last two. The four roles are:

1. Evidence as what justifies beliefs: it is the idea that our beliefs are justified when we proportion them to the evidence. This is an evidentialist thesis, regarded by many as a truism with respect to the justification of beliefs. And even those who reject the evidentialist thesis with respect to *prima facie* justification, normally adopt a non-defeasibility clause, that is, they accept that counter-evidence can defeat the justification of beliefs.[24]

2. Evidence as what rational people respect: here we have another quasi-truism that fits well with an evidentialist approach to justification of beliefs. Rational subjects respect their evidence. Just as 1, therefore, 2 seems to fit well with a psychologist's perspective on the nature of evidence.

3. Evidence as that which guides to truth: the notion of evidence here is of reliable guide. Smoke, for example, is a reliable signal or indicator of fire. Evidence, from this perspective, would then play a mediating role between our commitment to know reality and this reality to be known. Here the notion of evidence seems to be playing a role that fits better with an ontology of evidence in terms of facts, rather than psychological states, as in the case of the first two roles.

[23] Neta (2018) agrees with Goldman that E = K fails as a correct approach to evidence, but argues that Goldman's thesis would be incomplete, necessitating a condition of factivity. Neta's approach therefore seems to fit into what Goldman calls "other possible views as well, intermediate between E = K and NPJ."

[24] See Bergmann (1997) for discussion of how even externalists tend to adopt this type of clause.

4. Evidence as a neutral arbiter: evidence here is something public, objective, that adjudicates between competing theories or hypotheses. Or, as Kelly puts it, the "a kind of ultimate court of appeal" (2006). Here again, the notion of facts, as opposed to psychological states, seems to be more in line with this role played by the concept of evidence.

So while the first two roles seem to involve a notion of evidence in terms of mental states, the latter two seem to be evidence in terms of facts independent of what goes on in our individual mental lives. If this is correct, then there seems to be some tension between the different roles that the notion of evidence seems to play.[25]

An attempt to alleviate this tension is offered by Rysiew (2011), based on what Thomas Reid wrote about the nature of the evidence. According to Rysiew, Reid's position on the nature of the evidence offers an original combination of internalist and externalist aspects that seems capable of bringing unity to the different roles that the notion of evidence is expected to play. Reid's pluralism seems to be characterized by the combination of the ideas that "justifiedness of a belief is solely a function of one's evidence," and that the subject must "be in possession of a 'sound understanding' and in roughly the kind of world we take ourselves to be in" (222). The first part of the quotation would, given its evidentialist character, provide the resources necessary for the performance of the first two roles. The second part, because of its reliabilism or proper functionalism, would provide the resources to connect what is taken as evidence by the subject to the world, thus satisfying the requirement of factivity demanded by the other roles.

Another attempt to address this tension is offered by Mitova (2014), who argues, based on certain considerations of metaethics, for a position on the nature of evidence in which evidence is taken to be propositional, factive, and psychological. This combination of the three monistic conceptions of evidence presented so far is called by her "truthy psychologism." With this thesis about the nature of evidence, Mitova seeks to ensure that evidence is something capable of entering into logical and probabilistic relationships (the

[25] As far as Reid's position is concerned with respect to the nature of the evidence, Rysiew notes that "on Reid's rather liberal view evidence includes such varied things as propositions, states of consciousness, perceptual experiences, memorial seemings, statements, others' sayings and gestures, the rings on a tree, smoke on the horizon, the distinctive coloring of some type of bird, and so on" (220).

propositional component), can genuinely favor beliefs (the factivity component), and is in a position to favor beliefs (the psychological component). These three properties of Mitova's conception of evidence would, in her view, secure accommodation by her truthy psychologism of the four evidential roles we have just seen. The psychologism of this approach would allow the accommodation of the first two roles, and the condition of factivity would allow the accommodation of the last two. It would, indeed, be the possibility of this accommodation the main motivation for Mitova's truthy psychologism. The difficulties of psychologism in playing the roles that require a secure connection with the world would be circumvented in this proposal by the idea that it is "the veridical nature of certain mental state tokens is what suits those states for being evidence" (2014 , 19). By being formulated in psychological terms, this notion of evidence could play the roles of justifying beliefs and being what rational people seek; by having a formulation in terms of true propositions, it could play the roles of reliable indicator or neutral arbiter between theories.

EVIDENTIALISM

The idea that beliefs are justified when they are formed in response to the available evidence is intuitively plausible. Not surprisingly, it has played an important role in philosophical thought since at least the advent of modern epistemology, so much so that it is seen by many as a platitude (see Dougherty, 2011). According to John Locke, "He that believes without having any reason for believing [. . .] [does not seek] truth as he ought" (1690, book 4, chapter 17, §24). And "the mind, if it will proceed rationally, ought to examine all the grounds of probability, and see how they make more or less for or against any proposition, before it assents to or dissents from it" (1690, book 4, chapter 15, §5). More succinctly, David Hume states that "a wise man, therefore, proportions his belief to the evidence" (1748, §10). And even Hume's archrival Thomas Reid seems to agree: "To believe without evidence," says Reid, "is a weakness which every man is concerned to avoid, and which every man wishes to avoid" (1788, Chapter 2.20). And "all good evidence is commonly called reasonable evidence, and very justly, because it ought to govern our belief as reasonable creatures" (1788, Chapter 2.20).

The thesis enunciated by these philosophers resembles in large measure what is now known, especially through the work of Earl Conee and Richard Feldman, as evidentialism. As Andrew Moon (2012b) notes, virtually all contemporary epistemologists who identify themselves as internalists are evidentialists. However, there are also evidentialist externalists, such as Timothy Williamson (2000, 207-8), as well as attempts to construct hybrid models, or reliabilist evidentialism, to use Comesaña's expression (see Alston, 1988, and Comesaña, 2010). And even externalists who reject the thesis that justification of beliefs depends on their appropriate response to the available evidence include a no-defeat clause, *i.e.*, that the subject cannot have counter-evidence for the belief in question (see Bergmann, 1997).[26]

[26] Evidentialism is fundamentally a thesis about justification and, in a derivative way, about knowledge. Evidentialists are therefore generally concerned with providing necessary and sufficient conditions for rational, reasonable, or justified belief formation, not with providing necessary and sufficient conditions for warrant, that is, what, in addition to true belief constitutes knowledge. Moreover, evidentialism is not a theory about the morality of belief formation. The person who, in the stage of terminal illness, forms, contrary to the evidence, the belief that she will recover, may be doing something that, in practical terms, is quite appropriate – and even beneficial to her health – but that would be, in purely epistemic terms, rather inadequate (see

According to Kevin McCain, evidentialism is roughly the thesis that "facts about what a person is justified in believing supervene upon facts about the evidence that she has" (2014b, 1). A more precise formulation of the evidentialist thesis requires that one distinguishes between propositional justification and doxastic justification. The first concerns the existence of justification to believe. The second concerns the justification of the belief. In the evidentialist model, to say that a person has propositional justification for a belief is to say that she has evidence such that she is justified in taking a doxastic attitude toward *p*. An influential formulation of this idea by Conee and Feldman says that:

> (EJ) Doxastic attitude *D* toward proposition *p* is epistemically justified for *S* at *t* if and only if having *D* toward *p* fits the evidence *S* has at *t*. (Feldman and Conee 1985, 15).

However, it is possible that the person who has justification for believing a particular proposition will come to believe it for reasons other than those that propositionally justify the belief. Just as it is possible, ethically or morally speaking, to do the right thing, but for the wrong reasons, it is possible, analogously, to believe what the evidence justifies, but for other reasons. Thus, a complete theory of justification must also specify the conditions under which beliefs are based on evidence, which is what Conee and Feldman call a condition of well-foundedness (WF) for justification. Whatever the formulation of WF, it must satisfy the following schema:

> (WF) *S*'s doxastic attitude *D* at *t* toward a proposition *p* is well-founded if and only if
>
> *S* has *D* toward *p* on the basis of some body of evidence *e*, such that
>
> (a) *S* has *e* as evidence at *t*;
>
> (b) having *D* toward *p* fits *e*; and
>
> (c) There is no more inclusive body of evidence *e*' had by *S* at *t* such that having *D* toward *p* does not fit *e*'. [McCain, 2014b, 3]

Feldman, 2003, 41-45, for discussion of this distinction in light of W.K. Clifford's thesis that it is always wrong to believe in something without adequate evidence).

The formulation of a complete evidentialist thesis requires the development of the details of EJ and WF: (a) What is the nature of the evidence? (b) Under what conditions is it possible to say that someone *has* evidence? (c) Under what conditions does a belief *fit* the evidence? (d) What is the correct formulation of well-foundedness? In addition to these details, however, a successful defense of the evidentialist thesis involves the presentation of satisfactory answers to specific problems that have been raised against evidentialism, such as the justification of *a priori* beliefs, of stored beliefs, of forgotten beliefs, and of beliefs about the future, as well as the problem with forgotten defeaters. Let us see, first, how evidentialists have provided the details (a) – (d). Then, we will see what they have said in response to these problems.

Filling the Details of JE and BF

With respect to (a), as we have seen in the first chapter, there is disagreement among evidentialists as to whether evidence consists of mental states (see Conee and Feldman, 1985, 2004, 2008, McCain, 2014b), of the propositional content of mental states (Dougherty, 2011) or of propositions (Williamson, 2000).

With respect to (b), there are essentially three approaches to possession of evidence: there are inclusive approaches, which take as evidence all information stored in memory; restrictive approaches that limit the possession of evidence to what the subject is thinking or aware of at the time; and moderate approaches, which seek to find a middle ground between these two positions.

For a better understanding of these different positions, it is important to distinguish, as Feldman (2004, 226) does, total possible evidence (TPE) from total evidence (TE). TPE consists of "all and only the information the person has 'stored in his mind' at the time" (226), be it conscious or unconscious, recoverable or irrecoverable. It is excluded from TPE information that has been completely forgotten or of which the subject has never been aware of. TE, on the other hand, is a subset of TPE that excludes that which is part of TPE that has no justification.

An inclusive approach to possession of evidence says that TPE is equivalent to TE. Thus, even deeply stored memories, which would only be brought to consciousness through years of practice or psychotherapy, would be counted as evidence.

At the other end of the spectrum, there is the restrictive approach, advocated by Feldman (2004), according to which only what is being thought in the present moment is considered available evidence. One motivation for restrictive approaches is the possibility of escaping the New Evil Demon Problem (NEDP) for internalists put forward by Andrew Moon.[27] There are those like Kevin McCain (2014b, chapter 3), however, who consider inclusive approaches to be overly permissive and restrictive approaches as leaving out memories or beliefs that should be considered as available evidence. As a result, a moderate approach is proposed, according to which evidence is available to the subject if he is currently aware of the evidence or is willing to bring it to mind by reflecting on the truth or falsity of the proposition.

And under what conditions (c) does a belief fit certain evidence? That is, what conditions need to be fulfilled in order for the available evidence to support epistemically or evidentialy the belief in a particular proposition? There is a number of possible answers here. A good place to start is with the idea of logical entailment: *p* fits *S*'s evidence at *t* iff *S*'s evidence at *t* entails *p* (McCain, 2014b, 57). Two decisive problems with this approach to evidential support is that it excludes situations where support is inductive and includes propositions that *S* cannot even grasp but are implied by her evidence.

[27] The NEDP has emerged in the literature as a problem for externalist theories of justification. According to the NEDP, there is a possible world in which S* forms the same beliefs as S in the actual world, with the difference that S* experiences are caused by an evil demon. The predominant intuition here is that S's justified beliefs are the same as S*'s. However, S*'s beliefs are not formed reliably, and therefore, contrary to the prevailing intuition, S*'s beliefs would not be justified from an externalist perspective. Andrew Moon (2012a), however, sought to show that the problem also afflicts most internalist theories. According to Moon, there is a possible world in which an evil demon eliminates S*'s non-occurrent or unconscious. If S and S* have the same justified beliefs, then non-occurrent beliefs would be irrelevant for justification. If Moon's argument succeeds, internalists would have to take a restrictive position on the possession of evidence. See McCain (2014a, 2014b) for response to Moon's internalist version of the NEDP and Moon (2014, 2018) for answers to McCain's attempts to show that Moon's initial criticism (2012a) fail. What is at stake in this debate is whether the inclusive and moderate conceptions of possession of evidence are feasible. If Moon's arguments succeed, only restrictive forms of possession of evidence are feasible.

Alternatively, the condition could be formulated in probabilistic terms. However, while the alternative formulation of the fit condition in terms of what makes a proposition more probable is capable of solving the first problem, the problem of incapacity to grasp certain propositions remains, but now with respect to the probabilistic epistemic support they receive from evidence.[28]

Other possibilities, which have been shown to be much more plausible than the previous options, are those that use the notions of non-doxastic appearances and best explanation. Non-doxical appearances consist of mental states with propositional content that produce a sense that something is true. According to Phenomenal Conservatism (PC), if it seems to *S* that *p*, then, in the absence of defeaters, *S* has some degree of justification to believe that *p* (Huemer, 2007, 30). PC would provide, according to its proponents, a superior response vis-à-vis the alternatives of dealing with the skeptical threat from an internalist perspective, and the rejection of PC is claimed to be self-refuting (Huemer 2007; see DePoe, 2011, among others, for criticism to the self-refutation thesis).[29] Several problems have been pointed out with respect to this alternative, however. Among them: (i) non-doxastic appearances would not be necessary for justification, as in the case of one having strong reasons to believe in something (a mathematical proof, for example), although there is no appearance of this being true (Conee and Feldman 2008, McCain, 2014b); (ii) the problem of cognitive penetrability of perception, which poses difficulties for the sufficiency of the principle advocated by PC for epistemic justification (see, for example, Tucker, 2014, 12-16); (iii) the difficulty of incorporating perceptual appearances into a Bayesian approach (White, 2006;

[28] Another influential proposal regarding epistemic support, which sought to overcome the deficiencies of the deductive and probabilistic approaches, is that of Roderick Chisholm (1977). Chisholm proposed that principles regarding epistemic fit could be obtained from what we take to be known. The main problem with Chisholm's proposal is that these principles are not based on more fundamental and unifying principles, which leaves room for criticism that the choice of these principles is arbitrary. As Conee and Feldman put it, "it is difficult to resist the thought that these are principles designed to ratify the beliefs that Chisholm thought were justified" (2008, 97). Byerly proposes that there are two other approaches to evidence support that have proved impractical: that of Fred Dretske, who uses subjunctive conditionals, and that of C.I. Lewis, with evidential support in terms of consistency with the subject's evidence (see Byerly, 2014, note 19).

[29] A variety of PC, which takes inclinations to believe to be what gives justification to some degree in the absence of defeaters, is Richard Swinburne's credulism (name coming from the principle of credulity) (1979, 2018). Another variety is dogmatism, by James Pryor (2000) and Chris Tucker (2010), among others, the latter more concerned with justification of perceptual beliefs, while PC and credulism seek to be more general theses on justification.

Tucker, 2014, 16-20); and, (iv) if PC is true, absurd propositions (such as the morality of committing terrorist acts) could be non-inferentially justified (see Tooley, 2013; Littlejohn, 2011; and Huemer, 2013a, for answers to this objection; and Huemer, 2013a, and 2013b, for responses to others objections).

Another possibility, which has become popular among evidentialists (among the proponents of this position are Harman, 1973, Moser, 1989, Conee and Feldman, 2008, McCain, 2014, and Poston, 2014) is formulated in terms of best explanation. An initial formulation says that *p* fits S's evidence, *e*, at *t* iff *p* is part of the best explanation available to *S* at *t* of why *S* has *e* (McCain, 2014b, 63). This formulation, however, did not resist counterexamples such as those of Keith Lehrer (1974) and Alvin Goldman (2011) involving logical entailments. Since logical entailments are not explanatory, an adequate formulation of epistemic support in explanatory terms needs to take into account logical entailments of best explanations. With this in mind, McCain presents the following formulation of evidential fit:

> (EF) *p* fits *S*'s evidence, *e*, at *t* iff *p* is part of the best available explanation available to *S* at *t* for why *S* has *e* or *p* is available to *S* as a logical consequence of the best explanation available to *S* at *t* for why *S* has *e* (2014b, 63).

This formulation, however, is also deficient. According to T. Ryan Byerly (2013) and Byerly and Kraig Martin (2015), McCain's explanatory formulation of fit is incapable of accommodating justified beliefs of propositions about the future. McCain (2015) believes it is possible to salvage the evidentialist explanationism of this objection by substituting the "logical consequence" of the second formulation for "explanatory consequence" (Byerly and Martin persist with criticisms of explanationism in (2016) and McCain persists in his defense in 2017)).

The objection to the justification of beliefs about the future on explanationism is an objection to the *necessity* of evidentialist explanationist conditions for justification. Byerly and Martin (2016), however, also object to the *sufficiency* of the explanationist thesis, in cases where, despite the best explanation available to *S* being very good, the correct explanation is by no means available. One case that is presented to illustrate this is about a detective who examined what he considers to be only half of the available evidence, and at

that point a hypothesis emerges as being, not only good, but the best available. In this scenario, given explanationism, the detective would be justified in believing *h*. However, since there are good reasons to believe that the correct hypothesis may not be available, the detective would not be justified in believing *h* (see responses to these objections in McCain 2014b and 2017).

McCain's explanationism draws on some notions that need to be briefly clarified. The first is availability. Evidential explanatory support requires that evidence be available for *S* at *t*. McCain formulates availability in terms of *S*'s disposition to have an appearance that *p* is part of the best answer to the question 'Why does *S* have *e*?' based on reflection only"(McCain, 2014b, 66). Therefore, it is not necessary for *S* to be aware at the highest level of the evidential support in order to be justified, not even to have an appearance, but only that *S* is disposed, by virtue of her background beliefs, to have an appearance that connects *p* to the evidence (see McCain 2014b, 66-67 and 122). This conception of evidential support would make it plausible that even non-reflective children and adults have justifying evidence.

The second notion is that of total evidence. Like Poston (2014), McCain takes *e*, that is, evidence in his formulations of conditions of justification, as consisting of the total evidence of subject *S*, rather than a sub-set of her total evidence. If we focus only on a subset of *e*, we run the risk of ignoring the potential impact of defeating evidence. This is a crucial detail, for many people have difficulty taking the evidentialist thesis to be a plausible view of justification due to lack of attention to this detail.

As we saw earlier, a complete theory of justification must specify, in addition to the conditions for propositional justification, the conditions under which beliefs are based on evidence, that is, the condition (d) of well-foundedness (WF) for justification. There are essentially two possibilities here: to formulate WF in doxastic terms (*i.e.*, to say that a belief is well-formed is to say that we have meta-beliefs that support that conclusion) or in causal terms (*i.e.*, to say that a belief is well-founded is to say that it was formed "*because* of the evidence *e*"). The first option faces the serious problem that we rarely form meta-beliefs and therefore the adoption of the doxastic approach would imply that a large number of beliefs we take to be justified (and virtually all beliefs of children and non-reflective adults) would not be doxastically justified. The traditional problems with the second option are those of

causal deviation and overdetermination. The more sophisticated formulation of WF currently seeking to escape these problems is McCain's interventionist causal approach (2014b, chapter 5).

Three alternatives to conservatism and explanationism that deserve to be mentioned, albeit briefly, are direct acquaintance evidentialism (DAE), dispositional evidentialism, and probabilistic evidentialism. Direct acquaintance evidentialists argue that we can have direct contact with states, properties, or facts that are relevant to the truth of what is "given" to us by our experiences. This direct contact or awareness of what is "given" is a genuine relationship, since we cannot have direct contact with something that does not exist. For proponents of DAE, "all knowledge or justified belief ultimately depends on a foundation of knowledge or justified belief acquired by acquaintance" (Hasan, 2014; see also Moser, 1989, Fumerton 1995, McGrew 1995, Fales 1996, Bonjour, 2003, Hasan, 2013, and DePoe, n/a).

The second alternative, proposed by Byerly (2014), emerges from the idea that evidential support is a matter of being disposed to adopt a certain doxastic attitude toward *p* in light of *S*'s total evidence. As we have seen, McCain's explanationism also relies on the notion of dispositions. The crucial difference between Byerly's and McCain's proposals is that, unlike McCain's, Byerly's does not avail itself of the notions of appearance and best explanation.

Finally, there are different possibilities for formulating evidential support in probabilistic terms, with the most popular today being the Bayesian formulation. For Bayesians, support is a matter of subjective probability, with degrees of belief being rationally formed when they obey certain patterns of probabilistic coherence. Minimally, Bayesians say that degrees of belief are justified or rational insofar as the probabilities designated by *S* follow Kolmogorov's axioms and a conditionalization principle (see, for example, Talbott, 2008).

Problems and Responses

(i) Justification of the A Priori

A recurring claim against the evidentialist thesis is that *a priori* beliefs – such as logical and mathematical truths – are not formed on the basis of evidence. And, it is alleged, even if there was evidence involved in the formation of such beliefs, it would not be sufficient for

justification or warrant (that which distinguishes knowledge from mere true belief). Alvin Plantinga (1993, 188-193), for example, notes that *a priori* beliefs are often accompanied by images and sensations, but that they are fragmentary and vague and not constituting evidence. He recognizes that in addition to the sensory component, *a priori* beliefs involve inclinations or impulses to believe, or perhaps attractiveness or sense of inevitability toward certain propositions, and that perhaps this constitutes evidence. He argues, however, that inclinations to believe may, due to cognitive malfunction, not originate in response to the evidence, as in the case of the paranoid person who has a strong inclination to believe – irrationally – that her colleagues are plotting against her. Thus, even if inclinations or impulses could be considered evidence with potential justification for *a priori* beliefs, additional conditions, unrelated to evidentialism, would have to be added in order to construct a complete picture of the warrant conditions for *a priori* beliefs (see Feldman, 2004, 64-7, for response to Plantinga's treatment of impulsional evidence) (see John Greco, 2011, 170, for additional skepticism regarding evidentialist treatments of the *a priori*).

Conee and Feldman (2011, 286; 2004, 66) respond by listing some possibilities as to the nature of the *a priori* evidence: (a) the evidence is the appearance of truth regarding the proposition; (b) propositions known *a priori* are evidence for themselves; (c) inclinations to believe that the subject has learned that are reliable through previous successes and through the acceptance of the subject's interlocutors of their assertions involving *a priori* propositions; (d) impressions of detection of the truth-makers of the propositions. Conee and Feldman conclude that "the suggestion that our only evidential bases for simple arithmetic beliefs are impulses to believe is extremely implausible" (2004, 66).

According to McCain's explanationist model (2014b, 70), justification *a priori* comes from the reflection and understanding of the proposition in question, which produces some experience involving the awareness of the proposition and of the relations of its conceptual components. And if the truth of the proposition is part of the best explanation of these experiences, then the proposition fits the evidence. McCain (2014b, 158), however, acknowledges that more work needs to be done. In particular, we need to explain why the justification we have for *a priori* beliefs is far stronger than that of other beliefs.

(ii) Stored Beliefs

Our beliefs, even justified beliefs that seem to constitute knowledge, are not always present in our thinking. Most of the time, we are not "actively endorsing" the contents of these beliefs (Frise, 2017). They are non-occurrent, they are stored. More clearly, when we sleep and are not dreaming, all our beliefs are stored and, supposedly, we do not lose knowledge (for example, that the principle of noncontradiction (PN) is true) only because we are not awake (Moon, 2012b). However, what is the evidence base of beliefs in these circumstances? (see Goldman, 2011).

One response that has predominated among evidentialists is the dispositional one: dispositions of a certain kind, such as those that manifest themselves in appropriate circumstances, would be considered justifying evidence. The evidence base for these beliefs, on Conee and Feldman's views (2011), would consist of dispositions to retrieve these beliefs from memory, to bring them to mind as known, not merely believed. These dispositions would be mental states and therefore evidence: "Whatever causes a disposition to recollect [a proposition] does create evidence for the content proposition from scratch. Having the disposition constitutes having some defeasible evidence for its content"(2011, 305).

For McCain, however, such a response would not be entirely appropriate, since, in addition to the dispositions to recollect those beliefs, the beliefs would also have to be based on the dispositions. It is not clear, however, that beliefs can be based, that is to say, caused by, dispositions. He proposes, then, that in the case, for example, of our stored belief in PN, the justifying evidence would not consist only in the disposition to bring PN to mind, but in additional mental states, such as "a disposition for recollecting [PN] in virtue of having stored the content of [PN] in a particular way" (McCain, 2014b, 148; see also Frise, 2016, and Moon, 2018).

On the other hand, Frise (2018) argues that if representationalism about beliefs is true, stored beliefs simply do not exist and, therefore, there is no problem of stored beliefs. On the other hand, if the dispositionalist theory of beliefs is true (which is what Frise defends), the justifying evidence of stored beliefs would come in the form of *S*'s affirmation of

> *q* in response to a reason to affirm *q* — a recollective experience supporting *q*, a feeling of familiarity regarding *q*, or some other relevant phenomenon that occurs while he thinks about *q.*

which suggests that

> even when [one] isn't thinking about *q*, [one] has a disposition to have reason to affirm q. After all, every time [one] thinks about *q*, he has such a reason. Other things being equal, this disposition justifies his belief (2018, 71).

(iii) Beliefs with Forgotten Evidence

The evidentialist argues that the justification of a belief depends on *S*'s appropriate response to the available evidence. But if the evidence to which the belief responded is no longer available because it was forgotten, *S* would no longer have justification for her belief. The evidentialist might perhaps readily accept this conclusion were it not for the fact that there are cases of beliefs with forgotten evidence that seem to be justified. For example: last year Silvia read in a reliable magazine about the beneficial health effects of carrots. She continues to believe this, but she no longer remembers the source of this information, and has not obtained any additional evidence, for or against it. Silvia's belief seems to be justified (Goldman, 1999).

There are several ways to respond to this problem. One is to say that, since the health benefits of eating vegetables are widely publicized, Silva has evidence for the health benefits of carrots that are non-occurrent or unconscious but that can be easily brought to consciousness (Conee and Feldman, 2004, 70). She also probably has "supporting evidence consisting in stored beliefs about the general reliability and accuracy of memory. She knows that she is generally right about this sort of thing" (Ibid). As Frise (2015, 3b) puts it, " you have reason to believe that you tend to form beliefs with good reason, so you have evidence that you originally had good reason for your belief and this supports the belief." Or, the justifying evidence would consist of the "conscious qualities of the recollection, such as its vivacity and her associated feeling of confidence" (Conee and Feldman, 2004, 70).

McCain (2015, 475) exemplifies the case of forgotten evidence with the belief that 1776 was the year of US independence. He proposes that the evidence in such cases would come from the ability to distinguish genuine memory from mere imagination; the reliability of the person over the years in correctly remembering historical dates; the greater coherence of this belief with its other beliefs than with their negations; and the existence of meta-memories about the belief in question, that is, the existence of a remembrance that one knows something – in this case, the belief that the date of independence is known, preceding the recollection of the date itself.

(iv) Forgotten Defeaters

A problem analogous to the previous one is that of the forgotten defeaters. The main difference is that while the previous problem resulted from the loss of justification due to loss of evidence, the problem here is of restoration of justification due to loss of the defeater.[30] The problem of forgotten defeaters, however, is wider – for it affects more theories of justification – than the problem of forgotten evidence. This is due to the fact that there are more theories that accept evidence as contributing to the loss of justification (through non-defeat conditions, see, for example, Bergmann 1997) than as contributing to the justification of beliefs. A common response to this problem (see Frise, 2015, for discussion of literature) is to say that when one forgets the defeater, this leads to the loss of the effect of the defeater in one's noetic system, thus resulting in the restoration of the justification of the defeated belief. The problem with this answer is that many philosophers believe that memory only preserves (*i.e.*, it cannot create) justification.

(v) Beliefs about the Future

As mentioned previously, another objection to evidentialism, and to explanationist evidentialism in particular, is its supposed difficulty in accommodating beliefs about the future. A counterexample presented by T. Ryan Byerly (2013) showed that McCain's initial formulation of explanationism was inadequate. According to the counterexample, *S* gives

[30] Another difference is that, in the case of forgotten defeaters, it is possible that beliefs that have never been justified become justified. For example: Ana never had reason to believe that there are boxes in the basement, but she believes it anyway. If Ana forgets that she had no reason for this belief, she would become justified, according to some theories (Frise, 2015, 3.ii).

the last shot in a last round of a golf game in which she has performed well. The hole is very close and the ball is going in the right direction and speed. *S* forms the belief that the ball will fall into the hole. McCain (2015) acknowledges that this is a case of justified belief and that, as formulated initially, his explanationist condition of justification is not capable of giving that result. For this to occur, McCain acknowledges that it is necessary to substitute logical consequence for explanatory consequence in the second component of his epistemic fit. Byerly and Kraig Martin (2016) acknowledge that McCain's reformulation deals satisfactorily with the problem of beliefs about the future.

COGNITIVE SCIENCE OF RELIGION

Why are humans so prone to form religious beliefs? This question has captured the attention of thinkers for millennia. Many scientists believe that major progress has been made in answering this question. They believe that scientific experiments conducted by developmental psychologists and cognitive scientists of religion with young children have made an important contribution to our understanding of why and how humans seem so prone to become religious believers. In what follows, we will explore in a little more detail the findings of CSR. First, we will see some of the experiments that have led CSR researchers to identify a natural disposition in children to believe in the existence in the afterlife, in spirits, in super-knowing agents, their propensity to reason teleologically, to embrace creationism and substance dualism, in addition to experimental findings that suggest that our minds are not blank slates filled by our experiences. We will then examine the current debate on the naturalness thesis, the thesis that humans are naturally disposed to become religious believers. Are religious ideas and beliefs really natural? If so, in what sense are they natural? Finally, we will explore three main accounts of why religious thinking is so ubiquitous, paving the way for the discussion, in the next chapter, of the the implications of cognitive science (of religion) for (religious) evidentialism.

EMPIRICAL FINDINGS

Cognitive science and its sub-discipline CSR are empirical sciences. Their findings result, for the most part, from experiments conducted with subjects under controlled conditions. We will see below some of the main findings from CSR and some of the experiments that have contributed to the development of the theoretical underpinnings of CSR. Let us begin, however, with some of the findings that have led to the conclusion that the human mind is not a blank slate filled with “pure” experiences, one of the central findings of cognitive science that has important reverberations in CSR and whose comprehension will be crucial for our assessment, in the next chapter, of implications of the findings of cognitive science for the debate over evidentialism about epistemic justification (and , correspondently, of CSR for religious evidentialism).

a. Not Tabula Rasas

Several findings from cognitive science provide a formidable challenge to traditional empiricism – roughly the view that our minds are like empty slates that are filled by our experiences. Rather, as several experiments conducted by developmental psychologists and cognitive scientists strongly indicate, we come to the world with a large number of predispositions and inclinations that inform our behavior and how we respond to the world. The problem with empiricism, writes Justin Barrett, is that

> It ignores that human minds have a considerable number of natural tendencies that allow them to solve problems important for their survival and life concerns. From birth, human minds acquire and handle some kinds of information more efficiently than others (2012, 8).

One example of these tendencies is the capacity infants have to imitate facial expressions. In an experiment reported by Meltzoff and Moore (1983), forty newborn infants ranging in age from 0.7 to 71 hours were presented with both a mouth-opening and a tongue-protrusion gesture by an adult (figure 1)[31]. The results showed that infants can imitate both gestures. The imitation of facial expressions requires highly specialized capacities – matching external visual information with motor coordination – which the infants are able to carry out even though they have never seen their own faces in mirrors.

[31] This image is just an illustration, not from the original experiment by Meltzoff and Moore.

Figure 1

In addition to the compeling evidence that newborns are able to imitate facial expressions, there is also robust evidence that they "are born with some information about the structure of face [that] guides the preference for facelike patterns found in newborn infants" (Morton and Johnson, 1991, 164). Infants seem to be born with the capacity to distinguish human faces as a particular kind of visual stimulus. They pay more attention to human faces and notice the differences between them. The reason for this, notes Pascal Boyer, is that

> The child is quickly building a database of relevant people. From a few days after birth, the child starts to build up 'files' for each of the persons she interacts with, by remembering not only their faces but also how she interacted with them (2001, 110).

But

> [babies] do not start by seeing lots of stuff in the world and noticing that some of them – people's faces – have common features. They start with a predisposition to pay special attention to facelike displays and to the differences between them (Ibid.).

In their experiments with newborns, Morton and Johnson tested infants'preferences for specific structures of faces. They presented twenty-four newborns (mean age of thirty-seven minutes) with three head-shaped white forms, two of which with black features of a human face (figure 2). Their findings suggest that newborns have innate preferences for certain structures of faces. After the infants fixed their eyes on the drawings, which were presented idependently and randomly, the experimenter started moving the drawing in an attempt to elicit the infant's corresponding turn to the drawing in the new positions. The procedure was repeated to each of the drawings of the heads. The experiment was video recorded and analyzed by independent observers. The researchers found that "Face" and "Scrambled" elicited a much greater following behavior than did "Blank," and that the the infants followed "Face" farther than they followed "Scrambled."

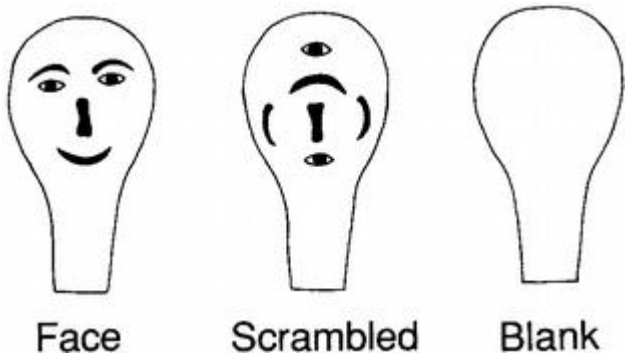


Figure 2

In order to make sure that infants are able to discriminate between intact faces and scrambled ones, Morton and Johnson presented them with figure 3, which included three alternative stimuli with face characteristics. Three groups of infants were tested – fourteen infants with mean ages of five weeks, fiteen infants with mean ages of ten weeks, and fifteen infants with mean ages of nineteen weeks. There first group didn't display significant preferences for any of the stimuli. The second group looked longer to "Face,"thus showing significant preference for "Face" compared to the other stimuli. This indicates that ten-week-olds favor not just facelike configurations, but prefer that the features of the face be located in their correct location. By contrast, and unexpectedly, the third group displayed significant differences among the four stimuli, with twelve of the fifteen infants preferring the other stimuli to "Face." A variation in the method of the experiment (instead of moving the stimuli, they were held fixed, while the infants were slowly moving on a rotating chair), however, showed that ten-week-olds favored "Face." The unexpected results found in the response given by the third group were, therefore, understood by the researchers as resulting from the testing method used. Overall, the results led the researchers to be very confident that infants prefer not just facelike configuration, but also that the features of the face be in correct locations.

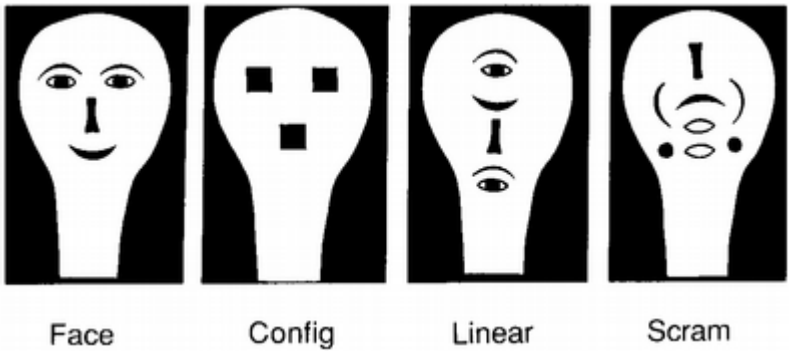


Figure 3

The conclusion that the human mind is not a *tabula rasa* can derived from additional studies. Simion, Regolin, and Bulf (2008), for instance, tested two-day-old babies' predisposition to attend to biological motion. They used a point-light animation, a traditional research tool used by cognitive scientists to assess subjects' responsiveness to biological and non-biological motion.[32] Prior to their research on newborn babies, studies aiming to assess the responsiveness of naive subjects[33] to biological motion were conducted with chickens. In one of these studies (Vallortigara et al., 2005), chickens were shown a point-light animation. The chickens prefered biological motion displays to any nonbiological ones regardless of the species of the animal depicted. The display (see figure 4) consisted of three frames with animation sequences: on the top, the biological motion stimulus (i.e., the walking hen); in the middle, the nonbiological motion stimulus consisting of random motion; and, on the bottom, an inverted biological motion display, consisting, for example, of an upside-down walking hen. This last frame was included to test the subjects' responsiveness to motion under condition of inversion effect.[34]

[32] The relevance of this research tool to investigate newborn's responsiveness to motion is explained by the researchers as follows: "The presence of animated motion in such displays is almost instantaneously detected by the visual system. Adults, for example, need as little as 100 ms to identify a point-light human walker. The locomotion of four-legged animals can also be promptly recognized by human observers, who can identify the different animals by their typical pattern of motion [. . .]. Several animal species are also able to discriminate and even to specifically respond to point-light displays depicting motion of conspecifics [. . .]" (809)

[33] Subjects that haven't been alive long enough to learn certain behaviors and responses to circumstances through experience and socialization.

[34] It has been frequently observed that biological motion perception is dramatically affected when the image is inverted.

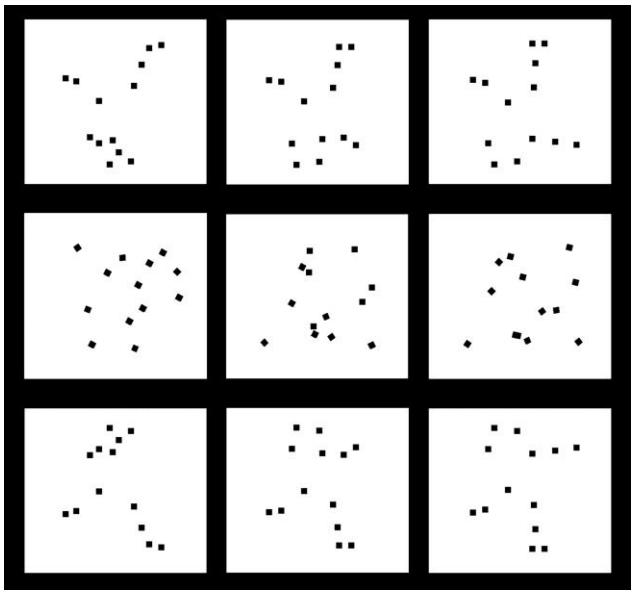


Figure 4

Simion et al. used the same point-light animation used in the chicken study in their study with two-day-old babies. They conducted three experiments aimed at detecting and measuring the newborns' different responses to the images. The results showed that the newborns were able to discriminate between the biological and the non-biological patterns of motion and prefered to look at the former. They also looked longer at the normal displays than inverted displays. Because of the young age of the subjects and the fact that the biological motion stimulus did not represent the shape of a human body or of any other familiar shape, the authors concluded that these results "cannot be accounted for on the basis of information acquired through learning"(812). This led the authors to hypothesize that these findings show that this detection of biological motion "is an intrinsic capacity of the visual system, which is presumably part of an evolutionarily ancient and nonspecies-specific system predisposing animals to preferentially attend to other animals" (809).

Researchers have also found that children of all ages assume that all members of a species have the same insides (see Boyer, 2001, 101, for discussion of the literature). In their experiments with children, Simons and Kial (1995) found that children lack expectations for what is inside animals and machines, but that they expect the insides of animals and machines to differ. If children are given a description of what is inside a certain animal – bones, muscles, and its internal organs – they infer that this will also be true of other animals

of the same species. Since children have little idea of what is inside the animals and were not apprised of any information that would lead them to think that what is true of the insides of one animal is also true of other animals, there isn't much in our experience to justify the conclusion that animal insides have the same insides. On the other hand, children are less certain that what is inside a specific sort of machine, such as a computer or a TV can be found in other computers or TVs. So, it is unlikely that children's expectations about the similarity of the insides of animals of the same species is the result of the way we classify and talk about them as species, for we also classify machines in similar ways – *e.g.*, we talk about this object being a computer and that object, though slightly different, also being a computer – but children don't see them as necessarily sharing the same internal components.

b. Omniscience

One of the areas of empirical research that has seen remarkable growth and gained prominence over the past three decades is the area that has been termed theory of mind (ToM). This area investigates children's understanding of other people's mental states – thoughts, perceptions, beliefs, emotions, desires, and so on. One particular set of experiments has revealed that children under four years old have difficulty understanding the notion of false belief. Wimmer and Perner (1983), for instance, found that children of that age tend to respond to stories involving a person re-entering a room after one of the objects of the room has been moved to another location by attributing knowledge to the person of the new location of the object, even though the person is in no position to know where the object is. Experiments like that, termed "false beliefs task," have been reproduced with variations and have produced similar results (see, among others, Rotman and Kelemen, 2012, for a review). One experiment involves an experimenter presenting children under five years old with a cracker box and then asking them what is inside the box. After the children answer "crackers," the experimenter shows them that there actually are rocks, not crackers, inside the box. The experimenter then asks the children how their Mom would answer the question "what is inside the box?" if their Moms entered the room at that moment. Most three-year olds answer "rocks," indicating that they believe that their mothers will have the right answer despite the appearance that there are crackers in the box. Tipically, unlike

children under four years old, five-year-olds will understand that their Moms may also conclude that there are crackers inside the box.

Barrett, Richert, and Driesenga (2001) have conducted a variation of the craker box version of the "false beliefs task" designed to explore children's views of God's knowledge. Their experiments proceed initially like regular craker box experiments: the experimenter presented children with cracker boxes containing rocks rather than crackers, the children answered "rocks," and the experimenter tells them that there actually are rocks inside it. But when it comes to the question about their Moms' belief, the experimenters included questions about a bear's, an ant's, a tree's, and God's beliefs about what is inside the box (experiment 1). Most four-year-olds told the experimenter that all these beings would believe that there were rocks in the box. And almost all five-and-six--year-olds answered "crackers." On the other hand, both groups of children answered "rocks" when asked what God would believe about the content of the box (with 53 participants, all four-year-olds and all but one five-year-old gave that answer).

The general conclusion to be drawn from "false beliefs tasks," according to Barrett and Richert, is that children "begin with a defaut assumption that beliefs are false and that beliefs are infallible and learn that beliefs can be wrong" (2003, 304). As we will see below, this view is known as the preparedness hypothesis. According to this view, children are cognitively "prepared" to believe certain things, including that all things in the natural world are created and that the creator or creators responsible for bringing natural entities about is/are super-powerfull and super-knowing. In CSR, the preparedness view was developed, as we will see below, as an alternative to the anthropomorphism hypothesis, which states, roughly, that children are incapable of conceiving God abstractly. Rather, according to this view, they anthropomorphise God, seeing him as a powerful and super-knowing human in the sky. Although the preparadness theory is one of the most widely accepted CSR theories today, there are cognitive scientists of religion, such as Rotmann and Kelemen (2012), who believe that the relevant data fits better with alternative approaches (in their case, with what they call the developmental approach, which we will discuss below as well). In any case, what is important to note here is that, regardless of which of these theories is the correct one, the evidence seems to point strongly in the direction of children developing at

some point the capacity to easily grasp the idea of a superknowing being. As Rotmann and Kelemen put it:

> In summary, it seems that children are able to reason accurately about superhuman minds very soon after they acquire the competence to think about human minds [. . .] it is also important to underscore the fact that five-year-olds are able to form concepts of religious agents that are qualitatively different from their concepts of other types of agents – a fact that cannot be accounted for by the anthropomorphism theory (2012, 15-16).

c. Dualism and Life after Death

Belief in immaterial beings is widespread throughout the world. Children seem to have a proclivity to believe in the existence of imaginary friends, and beings like ghosts and spirits can easily enter their imaginative universe. Belief in immaterial beings often goes unabated throughout adulthood. Adults, too, seem to have a proclivity to believe that the deceased may still be around, as spirits, and be able to communicate with the living. Belief in the existence of spirits, reincarnation, and so on, is part of the core teachings of many religious traditions and even agnostics and atheists may find themselves believing in the existence of ghost and spirits or at least occasionally finding themselves in circumstances in which they have uncanny experiences with what appear to be non-material beings.[35] If belief that the dead can still be around and perhaps communicate with us can be so widespread, perhaps we have a more general tendency to separate the world in terms of its material and immaterial elements. And perhaps this more general tendency has more specific components, with substance dualism and life after death being among the default assumptions of children. Several studies have been conducted to assess that possibility.

Philosophers have commonly used thought experiments about brain transplants to explore questions about personal identity and the nature of the human person. Popper and Eccles (1977), for instance, have pointed out that "the flawless transplantation of the brain, were it

[35] See Barrett (2012, 112-120), for instance, for a couple of stories about agnostics and atheists who found themselves in such circumstances.

possible, would amount to a transference of the mind, of the self" (117). This seems to be in fact the judgement of most adults. Experiments involving brain transplants, however, have reveled that children see the brain as responsible for only certain aspects of our mental lives, typically those involving deliberate mental work, such as in math problem solving, but not those involving physical action, memory, and related to identity and personality.

In four experiments with a total of almost two-hundred children with ages from four to ten, Carl Johnson (1990) assessed children's understanding of the subjective grounds of identity by examining their judgements about the consequences of hypothetical brain transplants. The children were aked questions about the effects of the transplant of the brain and other body parts, such as the mouth, the face, and the heart, from the child to another person and to non-human beings, such as pigs. The youngest children (ages five to seven) were able to imagine the consequences of transplants from themselves and a pig or a baby, but the findings suggest that children do not think that a brain transplant would lead to a complete transfer of the mental states of one person to another. He asked a series of questions about a post-transplant pig, such as, "With your brain, would this pig have memories of being a pig?" "With your brain, would this pig have memories of being a child?" 12% of the kindergartners claimed the pig would have attributes of a child and no attributes of a pig; 25% responded that the pig with the child's brain would acquire no child characteristics; 46% of first-graders thought that the pig with the new brain would have only child characteristics; 8% concluded that the pig would have only pig characteristics after the transplant. Although Johnson himself did not derive this conclusion in such a straight-forward way (see Gottfried et al, 1999, for discussion of Johnson's more limited conclusion), the children did not think that the pig would necessarily acquire the thoughts and memories of the person associated with that particular brain.

In another series of experiments conducted by Gottfried, Gelman, and Shultz (1999), children were told stories about transplants of the brain and other body parts between animals and asked questions about the transplant of these organs and parts designed to test the children's views about whether animals have essences and what it would consist in. In particular, the experimenters wanted to assess the children's views about whether all thoughts and memories would be transferred along with the organs and body parts and

whether they find those organs and parts essential for mental activities related to thought and memory. Seven-year-olds and younger children often indicated that the brain is important, but not essential for mental activity involved in thoughts and memories. Their findings challenged the idea that children think of the brain as a container – "a filing cabinet that holds thoughts, memories, ideas, and other mental products within it"(148), thus providing further confirmation of Johnson's findings (1990).

Lillard (1996) conducted five experiments designed to assess whether children think of pretending as a mental state. Previous research on pretend play had provided evidence that young children do not understand pretending as involving mental states. Lillard sought to assess whether those results hold under different experimental conditions. In one of the experiments, the children were shown photographs containing images of characters performing mental and physical processes (such as imagining, in the first case, and walking, in the second). Then the children were asked if the character needed his/her brain to perform that process. She found that children see pretending as consisting primarily of physical processes rather than mental processes. Pretense, in oher words, does not, in their view, involve the mind. This is significant in that it provides evidence that, for children, mental functions and psychobiological functions need not work together.

Paul Bloom offers the following assessment of the evidence coming from these and similar experiments:

> [Children] do not usually understand that the brain is needed for physical action, such as hopping and brushing your teeth, and they do not think the brain is needed for an activity like pretending to be a kangaroo. And if you tell these children a story in which a child's brain is successfully transplanted into the head of a pig, children agree that the pig would be as smart as a person, but they would think it would still keep the memories, personalities, and identity of the pig (2004, 200).

Kuhlmeier, Bloom, and Wynn (2004) conducted experiments with five-months-old infants aiming to evaluate whether they see human beings as material objects. Five-months-old

infants understand objects as solid and cohesive and capable of moving continuously through space. If they see human beings as material objects we should expect them to understand humans as solid and cohesive and under the constraints of continuous motion as well. Kuhlmeier et al. found that while five-month-old infants apply these constraints to inanimate blocks, they don't do the same to human beings. They showed the participants videos of humans disappearing between two barriers and then reappearing again. The lack of surprise on the part of the infants by the discontinuous motion of humans through the barriers indicate that "infants have two separate modes of construal: one for inanimate objects and another for humans" (2004, 95). Infants, in other words, do not seem to "readily view humans as material objects" (2004, 101).

The strongest evidence that children are Cartesian dualists come, however, from studies that show that children don't see death as constituting the end of one's mental life. In an experiment reported by Bering, Blasi, and Bjorklund (2005), Spanish children watched a puppet show about Baby Mouse. As the story goes, Baby Mouse goes for a walk in the woods and is hungry and upset with his brother. Baby Mouse ends up being eaten by an alligator and the experimenter makes sure the children understand that Baby Mouse is dead. The experimenter then proceeds to ask the children questions such as "do you think Baby Mouse is still hungry now?," "do you think Baby Mouse still wishes that he didn't have a brother now?" The children were more confident that Baby Mouse was no longer hungry than that he was no longer angry at his brother. In other words, after Baby Mouse's death, they were more certain that he no longer had physical properties than that he had mental properties.

Jesse Bering (2011; and 2004 and 2005 with his collaborators) and Paul Bloom (2004, 2011), among others, have argued that these and other findings show that children are substance dualists. As Bloom summarizes his views:

> [Children] are dualists in the sense that they naturally see the world as containing two distinc domains, what [psychologist Henry] Wellman calls 'physical objects and real events' and 'mental states and entities' – what I have described as bodies and souls (2004, 199).

Rottman and Kelemen (2012) agree that the empirical evidence supports the view that children's cognitive developmeny lead them to view humans as having immaterial souls, but disagree with Bering and Bloom about their understanding of "natural." While Bering's and Bloom's understanding of the naturalness of dualism suggest that such beliefs are "innate," Rottman and Kelemen believe that naturalness should be understood in developmental terms (2012, 23), as we will see below. Still, Rottman and Kelemen's conclusion about the strength of the evidence in support of the idea that dualism is our default assumption about the nature of the human person is similar to those of Bering, Bloom, and others:

> Dualistic thinking does develop at some point during early childhood, however, and at least in the context of reasoning about death, this may come about as early as four years of age (Bering, Blasi, & Bjorklund, 2005). When asked questions about the types of processes that continue after death, especially after hearing death described in a religious narrative, children exhibit dualism by believing that mental functions persist while bodily and psychobiological functions cease (2012, 23).

d. Teleology

A series of studies conducted by Deborah Kelemen and her collaborators has shed new light on something psychologists have been discussing since at least Jean Piaget: children's propensity to reason teleologically about natural entities. In the interviews he conducted with young children, Piaget often obtained purpose-based explanations for natural phenomena. Piaget found that when asked questions like "why it gets dark at night?" children would often give answers like ''so that people can go to bed'' (Piaget, 1972, 294). He concluded that this teleological proclivity only applied to children's thinking about natural entities and that they identified the purpose of things with human agency. Subsequent research has shown that Piaget was right about children's inclination to see natural kinds as existing for a purpose. But it has also shown that he was wrong in thinking that children's teleological thinking only applied to natural kinds and that the agency behing the existence of these kinds is human.

Several experiments have provided robust evidence that young children have a promiscuous tendency to view not only non-living natural phenomena (such as rocks, mountains, and storms) as existing for functions or purposes, but also plants and animals. When asked why rocks are pointy, children often give teleological answers such as “so that animals will not sit on them,” or “so that they can scratch their backs” (Kelemen, 1999a). But they also give teleological answers when asked why living entities exist. In a study Kelemen (1999b) designed to assess the scope of teleological thinking in pre-school children, the experimenter invited the children to play a game with the experimenter and two fictitious charecters in photographs: “Ben” and “Jane.” The characters were introduced as persons who “love to talk about different things but never ever agree with each other”. The children were told to listen to what Ben and Jane had to say about a certain subject and then point to the character that they thought was right. One of the items presented to them was the following:

> See this. This is a tiger.
>
> Ben says a tiger is made for something. It could be that it’s made for eating and walking and being seen at the zoo or it could be that it’s made for other things. But Ben is sure that a tiger is made for something and that’s why it’s here. Jane says that this is silly. A tiger isn’t made for anything. Even though it can eat and walk and be seen at the zoo, that’s not what it’s made for, they’re just things it can do or people can do with it. Jane is sure that a tiger can do many things but they aren’t what it’s made for and they aren’t why it’s here.
>
> Point to who you think is right. Ben who thinks a tiger is made for something or Jane who thinks that’s silly because a tiger isn’t made for anything (Kelemen, 1999b, 256).

The children were as likely to say that the teleological explanation was the correct one as to agree with the non-teleological answer. Given that American parents don’t seem very propense to teach their children that living entities such as tigers, birds, and trees owe their existence to a function or a purpose like “so that it can be seen at the zoo” or (in the case of

birds) “so that they can look pretty,” it seems that these results cannot be explained by cultural influence (see Barrett, 2012, 47, for discussion of this point).

The traditional view about human teleological thinking (which Kelemen calls “selective teleology”) argued that teleology is an innate mode of thinking applied selectively by children and adults to non-living entities. But the experiments conducted by Kelemen and her collaborators seem to show that teleological reasoning is not restricted to any particular category of phenomena. Rather than having selective teleological propensities, children seem to possess “promiscuous teleology” (to use another of Kelemen’s expressions), applying it to all kinds of entities.

Children seem to reason about both non-living and living entities in terms of functions and purposes, more so than adults, studies have unsurprisingly shown. But adults retain a strong propensity for teleological thinking. In fact, even highly-educated adults are prone to reason teleologically under time constraints, indicating that while the tendency to endorse purpose-based explanations may be attenuated with adulthood and education, it is rarely completely eliminated. In a study conducted with physical scientists at top-ranked U.S. universities, it was shown that those subjects displayed the tendency to reason teleologically when their information-processing resources were limited (Kelemen, Rottman, and Seston, 2013). Other studies have found similar results with respect to adults trained in other academic disciplines (e.g., Shtulman and Valcarcel, 2012). In sum, these studies show that while highly-educated adults may not display teleological biases under ordinary conditions, they tend to do so under highly speeded conditions.

Other studies have found that Roma adults without formal education (Casler and Kelemen, 2008) and Alzheimer’s patients have higher-level of teleological biases than other adults (Lombrozo, Kelemen, and Zaitchik, 2007). The study with Roma adults included both formally educated and non-formally educated Roma and the results were compared to another study with U.S. educated adults. Roman adults without formal schooling displayed stronger propensity for purpose-based thinking than the two other groups. The authors of the study concluded that these findings provided further evidence that purpose-based reasoning is not just a bias characterist of he early stages of human cognitive development that naturally disappears as one reaches a certain level of cognitive development. Rather,

they took these findings to indicate that a propensity for teleological reasoning represents a cognitive default that may persist throughout life. Further evidence for this thesis was obtained from the finding that Alzheimer patients display greater levels of bias toward purposeful explanation.

One common criticism of the promiscuous teleology hypothesis is that the evidence for a propensity for purpose-based explanations that reaches the level of cognitive default has come primarily from Western or Western-influenced cultures, with its Abrahamic theistic background. This has led researchers to seek new frontiers to test their hypotheses. China, as a non-western society that has not been under the cultural influence of Abrahamic theism, and in fact, that has been strongly influenced by an atheist worldview and that has only a small percentage of the population claiming to believe in God, has become a fertile ground for the testing of the promiscuous teleology hypothesis and other theories pertaining to the naturalness of religious belief. Several recent studies in China have lent support to the naturalness thesis and to the idea that the propensity to endorse purpose-based explanations is a universal cognitive default (Rottman et al., 2016, Hornbeck et al., 2017, Järnefelta et al., 2018). These studies, however, have also found that the lack of theistic cultural background contributes to attenuate that default teleological tendency.

e. Creationism

In addition to promiscuous teleological thinking, children also have a strong proclivity to be creationists and have difficulty accepting the theory of evolution. In an experiment reported by Margaret Evans (2001), 185 children from five to ten years of age and 92 mothers from both Christian fundamentalist and non-fundamentalist backgrounds were interviewed about their views on the origins of animate and inanimate entities and artifacts, and on their knowledge of natural history. The children were shown pictures of animate entities (sun bear, tuatara, human), inanimate entities (rock, crystal) and two artifacts (toy chair, doll) and asked “how do you think that the very first [one of those entities or artifacts] got here on earth?” (227). For each pictured item, the children were asked to what extent they agreed or disagreed with statements about the origins of the entity from the perspective of a Creationist (‘‘God made it and put it on earth’’), of an Artificialist (‘‘A person made it and put it on earth’’), of an Evolutionist (‘‘It changed from a different kind of animal that used to live

on earth''), in addition to spontaneous-generationist explanations (''It just appeared; It came out of the ground''). The parents were asked similar questions, but phrased as if their children were asking them the questions ("how do you think that the very first [one of those entities or artifacts] got here on earth?"). The answers given by the children (both from fundamentalism and non-fundamentalist parents) showed a creationist bias. While this might be expected from children from fundamentalist backgrounds, non-fundamentalist children surprisingly showed greater affinity for creationism than their parents. Criationism was favored over evolution by children of both groups and ages, though adeherence to creationism over evolution was particularly strong among the younger children. In sum, the findings from this and other studies have shown that children have difficulty accepting evolution and consistently favor creationist accounts of the origins of species over evolutionary ones.

In another set experiments, George Newman and his collaborators (2010) found that the groups tested – ranging from twelve months of age infants to six-year-old children – see order as created by agents, not impersonal causes. In the first experiment, the experimenters showed three-to-six-year-old children the drawing of a room and told them that that was Billy's room and that he went outside. Half of the children were then told that his sister Jane went into the room and changed his things. The other half were told that the wind blew through the window and changed his things. The children were then shown two test cards (figure 5). One card showed objects arranged orderly. The other card showed a disordered arrangement. The children were then asked "which one of these piles looks most like Julie (the wind) changed it?"

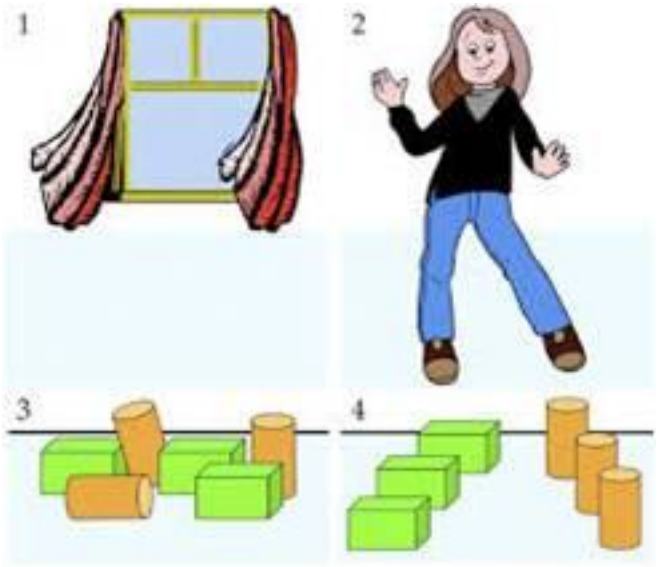


Figure 5

Newman at al. found that the children viewed Jane (the agent) as significantly more likely to make the ordered arrangement than was the wind. Eighty-eight percent of the responses attributed the disordered arrangement to the wind (the inanimate item), and twelve percent attributed the ordered arrangement to it. On the other hand, sixty-two percent of the responses attributed ordered arrangement to the agent, whereas thirty-eight percent attributed the disordered arrangement to her.

The second experiment tested infants belonging to two age groups: twelve-months-olds and seven-months-olds, totaling 48 infants. Both groups were presented with movies of two ordering and two disordering events. Fisrt the infants watched a ball changing a disordered set of blocks into an ordered arrangement (figure 6). The film begins with an ordered pile of blocks. Then a barrier moves in front of the blocks. A ball subsequently appears and moves behind the barrier. The barrier is then dropped and the blocks are disordered. This scenario is modified so that there is a combination of (2) initial disorder, the appearance of a ball, and then order; (3) initial order, the appearance of an agent, and then disorder; (4) initial disorder, the appearance of an agent, and then order.

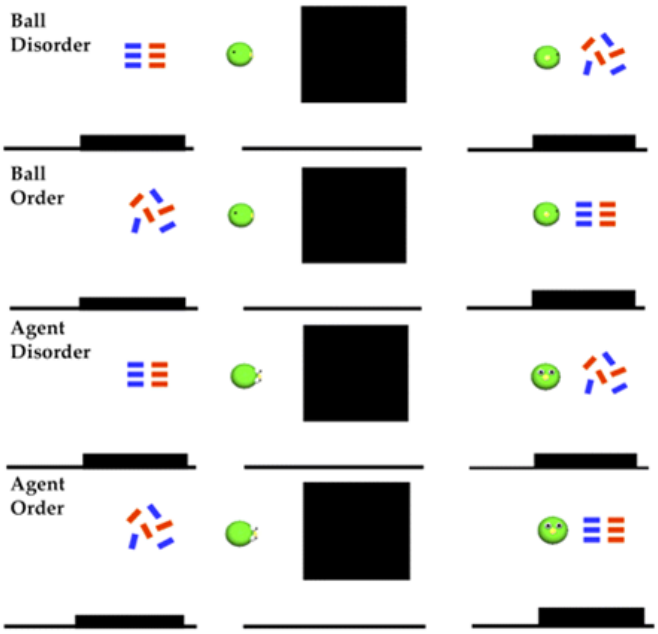


Figure 6

The experimenters measured the infants' surprise or lack thererof in each scenario by measuring their looking-time pattern – the longer they looked at the outcome, the more surprised they were supposed to be. The results showed that only the older group of infants was significantly surprised by a ball (the inanimate item) producing order. Unsurprisingly, none of the groups was surprised by the outcome of the agent producing order. These experiemtents reveal that

> by at least 12 mo of age, infants appreciate that agents are capable of creating order, whereas inanimate objects are not. Moreover, at 4 y of age, children readily identify that, whereas agents are capable of creating either order or disorder, inanimate forces (such as the wind) are capable of creating only disorder (Newman, 2010, 17144).

The conclusion the authors derive from these findings is that by twelve-months of age infants understand that "agents causally intervene on the world in fundamentally different ways from inanimate objects"(17140). By such an early age, we are already capable of understanding that while disorder can be created by both agents and inanimate objects, only agents can create order.

f. Spirits

Religion is normally associated with belief in supernatural agency. A religious person is not uncommonly – and perhaps rightly so – viewed as a person who believes that there are beings that are not part of the physical world – perhaps God, gods, spirits, angels, demons, and ghosts – but that can perhaps at least interact with it in some way or another. Many contributors to the field of CSR have suggested or even explored in detail and defended the idea that belief in supernatural agency can be traced to the human overattribution of intentionality and agency. And several pieces of evidence seem to point in this direction.

In the experiments conducted by Simion, Regolin, and Bulf (2008) with newborns that we discussed above, it was found that two-day-old babies already attend preferably to biological motion. The results showed that the newborns were able to discriminate between the biological and the non-biological patterns of motion and prefered to look at the former. In other words, human beings seem predisposed to detect agency. We are born with a default preference for focusing our attention on other agents. The preference of the newborns for biological motion cannot be accounted for in terms of a learning process – there is not enough time for learning those responses when you are only two-days-old! Rather, it must be accounted for in terms of, as the authors hypothesized, "an intrinsic capacity of the visual system, which is presumably part of an evolutionarily ancient and nonspecies-specific system predisposing animals to preferentially attend to other animals" (809).

In addition, as reported by Luo (2011), infants attribute agency to ambiguous stimuli. In two experiments, Luo sought to examine whether three-month-old infants would attribute agency to a self-propelled box when cues were given to infants that the box was an intentional agent. Luo found that the infants increased their attention when the box approached an object. They acted as though the box would maintain this goal and increased their attention accordingly. They interpreted the box's actions as goal-directed rather than as the consequence of impersonal stimuli. According to Luo, the experiments demonstrated that infants as young as three-months engage in intentional understanding about both humans and non-human agents by attributing goals to them. Luo hypothesized that this is

the result of “an innate psychological-reasoning system [that] may guide infants’ interpretation of all agents” (459).

One important finding from studies such as those conducted by Luo is that while infants attribute agency to inanimate objects, they expect the object to move only when contacted by another moving object (see Saxe et al., 2005, for na overview of this literature). In the case of Luo’s study, they attributed goal to the box only when the experimenter was present. Other studies have built on these findings and investigated infants’ expectations about the source of causal power. In an experiment conducted by Saxe and her collaborators (2005) with ten- and twelve-month-old infants, it was found that the infants expected a human hand, not an inanimate object, to be the primary cause of the object’s motion and that the infants can infer a hidden agent without direct perceptual evidence. The experimenters set up a stage with a wall in the middle (figure 7). A beanbag was thrown over the wall. Initially, the event was hidden, with the beanbag emerging already in motion. Adults perceive the beanbag as being thrown and the experimenters wanted to find whether the infants would share that perception. If so, the infants might expect to see a human hand on the side of the origin, but not on the opposite side (figures 7b and 7c). Controlled conditions (figure 7d and 7e) were also included, with a nonagent (a toy train – 7d – and and a puppet – 7e) replacing the hand.

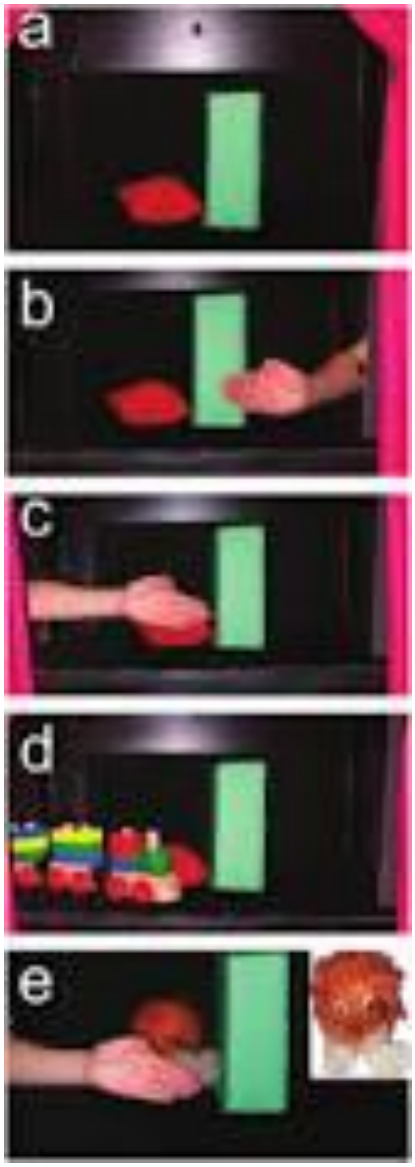


Figure 7

Forty infants participated in the study. Half of them participated in the hand test trials, and half in the train test trials. The researchers found that the infants looked longer at the human hand that emerged from the opposite side of the stage than at the hand that emerged from the same side as the beanbag, indicating that they expected the hand to be on the side where the beanbag originated. There was no difference between same-side and different-side trials when the toy train instead of the hand appeared, indicating that they didn't expect the train to be on the side where the beanbag originated and, therefore, that they didn't see the train as the cause of the motion of the beanbag. In another experiment the object entering the stage was a self-propelled puppet (figure 7e). The rationale for this

final test was to verify the infants' respose when there was an alternative explanation to the existence of a hidden hand. In this new trial, the infants did not look longer when the hand emerged from the side opposite the puppet than when the hand emerged from the same side, indicating that once they recognized that the object entering the stage was itself an agent, no hidden agency should be expected. Saxe et al. conclude from these finding that

> These results indicate that 12-month-old infants represent the invisible causal agent of an inanimate patient's motion and consider a person a more likely causal agent than a train. After a beanbag ''flew'' onto the stage from one side, infants showed less surprise when a hand suddenly appeared on the stage on that side than when a hand emerged from the other side; looking times were equal for a train that emerged on the same side as the beanbag and a train that emerged from the other side. These results suggest that the infants distinguish between hands and trains as potential causal agents. Furthermore, the interaction between conditions rules out the possibility that infants' looking time in the hand condition was governed by a simple spatial association (e.g., ''all moving things come from the left'') (998).[36]

These are only three examples of researches on infants' and children's views about human agency that seem to point to the existence of default cognitive mechanism in humans for agency detection. Consistent with Saxe's findings (2005, 2007), other studies have found that infants do not expect agents to necessarily be visible or resemble humans (see Barrett, 2012, 29ff, for discussion and description of experiments). These and other findings seem to suggest that humans are, as Barrett put it (2012, 34), "eager detectors of agents." In fact, infants and children are prone to identify a wide variety of objects and entities, such as shadows and strange lights, as agents. Even adults, in certain circunstances, may find themselves easily identifying certain objects and entities as agents and to take certain noises as indicative of the presence of other agents. We not uncommonly perceive certain objects

[36] Similar experiments were conducted by Saxe et al., 2007, with the same results.

or entities, such as clouds, to resemble human faces or other characteristics of agents. This hás led several authors to suggested that we have a specific faculty or a combination of faculties specifically devoted to the detection of agency and that this faculty or faculties are hypersensitive – they are so prone to detect agency that often lead us to perceive agency where there is none. This faculty has been termed hypersensitive agency detector device (HADD) and has been embraced by many as a central factor in the emergency of religion as it would have conferred evolutionary advantage to those overdetecting agency in comparisson with those without this propensity.

THE NATURALNESS THESIS

"Many cognitive scientists," note Konika Banerjee and Paul Bloom, "see the universality and pervasiveness of religious belief as suggesting that it is a natural feature of evolved human psychology" (2013, 7). This is, in a nutshell, the natural thesis. Religious belief is, cognitively speaking, said to be natural. But in what sense is religious belief natural? Is it natural in the same way that breathing is natural for human beings? Or would religious belief be natural more in the sense that walking is natural? Or perhaps is it natural in the sense that writing is natural for humans who have learned how to write?

Robert McCauley (2011) has developed a helpful terminology. He distinguishes two types of naturalness for the purposes of understanding the naturalness of religion: practiced naturalness and maturational naturalness. The first consists of skills that have been invented by humans at a particular moment of history, such as writing and riding bycicles. These skills are not natural in the sense that we are born with them, but in the sense that once we undergo the learning process necessary to master them above a certain threshold of competence, they become "familiar, automatic, and unconscious," and "eventually begin to feel natural" (24). The origin of these skills is then cultural, but once they are properly nurtured, they become second nature.

The second type of naturalness consists of skills that children in every culture will develop, such as chewing, walking, and using one's native language. All humans with normal development will eventually learn to chew, to walk, and to speak her native language, regardless of the cultural milieu one is part of. As the name suggests, maturational

naturalness does not involve capability to use the skills in question from the moment of birth. While practiced naturalness has its roots in culture, being at some point in time invented by someone, maturationally natural skills are universal and inevitable, or nearly so.[37]

Notice that neither practiced nor maturational naturalness involves the sort of naturalness found in, for example, breathing. We naturally breath from the moment we emerge from intrauterine life (though some of us need a gentle slap from the doctor!), but we don't chew, walk, or speak our native language from that moment. Both are natural, however, in the sense inevitability. And this brings to mind the notion of innateness. If religion is natural in the sense that it is like chewing, walking, or speaking one's native language, is it innate? Many cognitive scientists who embrace the notion of naturalness (universality plus near inevitability) eschew the notion of innatness when speaking of the naturalness of religion. One of them is Justin Barrett. Barrett (2012) breaks from McCauley's terminology, preferring to speak of natural traits or nurture (we naturally become language users or walkers) and expert traits or expertise (we acquire expersite in writing and – some of us – in dancing ballet, which then becomes automatic, "second nature"). The advantage of this terminology, according to Barrett, is that it allows for more gradation in terms of naturalness: "Being able to add 1 + 1 might be fully natural, and adding larger sums might be mostly natural, but doing calculus is very unnatural" (19). But natural traits are not necessarily "built into our biology from birth, or hardwired into our brains" (ibid). They may constitute biological dispositions, alright, but that doesn't mean their development does not require the right sort of environment. "What we can more sensibly say," writes Barrett, "is that given a certain kind of biological endownment and the ordinary sort of world we are typically born into, we will typically develop certain properties and attributes" (19). And this combination of normal biology plus typical environment is what is involved when CSR theorist like Barrett speak of natural traits and, in particular, of religion being natural. Barrett (and others), on

[37] Both types of naturalness belong to the broader category of intuitive cognition, as opposed to reflective cognition. Intuitiveness or absence of reflection, however, is not sufficient for naturalness in those two senses. There are certainly thoughts or actions that are non-reflective or intuitive in the way they are originated but which are not natural.

the basis of the findings of CSR (some of which we discussed above), believe that belief in supernatural agents, including gods of some sort or other, and maybe God, are natural in this sense of biology plus the typical environment where humans develop. For these authors, certain religious beliefs are not innate. They are not like breathing. But they are not something dependent on cultural particularities. They result from early-developing cognitive systems that "make belief in some kind of god almost inevitable" (20). They make children quite susceptible to believinge in gods because "gods occupy a sweet spot in their natural way of thinking: gods are readily and easily accommodated by children's minds and fill some naturally occurring conceptual gaps rather nicely" (20).

Barrett's formula "biology + environment" is a sort of middle way between Jesse Bering's (e.g., 2006) (and perhaps Boyer's (2001)) more strict focus on the biological component, on the one hand, and Bloom's (2007, Banerjee and Bloom, 2013) and Rottman and Kelemen's (2012) acceptance, on the other hand, of the role that culture may play in nourishing the biological traits that incline us to embrace certain religious ideas. Unlike Scott Atran, Barrett, Bloom, Boyer (among others) who are proponents of the by-product view of religious development (i.e., religion is an evolutionary by-product, accident, or spandrel of certain mechanisms, such as agency detection, that have adaptive value) Bering is a proponent of the view that religious ideas are evolutionary adaptations (i.e., religion was directly selected in our evolutionary history because of its influence on altruistic thought and behavior that proved to be adaptative). This leads Bering to predict that certain religious ideas, such as of the afterlife, are inevitable. Bloom, on the other hand, claims that common-sense dualism and agency detection will lead us to be easily receptible to religious notions, though not make religious belief inevitable. Bloom compares religion to language: "Like language, religion is universal. All societies have at least one language; all societies have at least one religion." But,

> Also like language, religion is not present at birth. It develops instead through immersion in a social environment. The specific language or religion that a child develops is determined by the culture in which the child is raised, not by genes or the physical environment (2007, 148, 149).

This leads Bloom (and Banerjee) to claim that, while natural – in that we have early-emerging cognitive biases that give rise to religious ideas – its naturalness consists in receptivity, not generativity (Banerjee and Bloom, 2013). We are naturally receptive to religious ideas, and once we acquire them, they easily become part of our cognitive and doxastic repertoire. But religion is not easily generated in the absence of cultural stimuli. Would Tarzan believe in God in the absence of certain cultural milieu that could provide him with some religious conceptual background? Their answer is no. He could easily embrace and preserve religious notions if they were at least to some degree in circulation in his environment, but not in absence of them. Those early-emerging biases are "the seeds from which religion grows" (2007, 150) but the seed needs good soil and watering.

Another postion in this debate is that of Rottman and Kelemen, with their emphasis on developmental constraints on the naturalness of religion. They agree with the vast majority of CSR researchers that naturalness is an appropriate term to express what is involved in the process through which humans embrace religious ideas. However, they disagree with researchers who take a stonger view of the thesis that religion is natural. Some authors, such as Bering, take religion to be somehow innate and inevitable because of its evolutionary origins. Others, such as Barrett and McCauley, see religious belief as being "maturationally natural" or a "natural trait," coming very close to being inevitable if the biological component is combined with environmental preconditions. Still others, such as Banerjee and Bloom accept a slightly more moderate understanding of the naturalness thesis: they see religion as emerging from built-in biases, but these biases need the presence of certain social preconditions in order to produce religious beliefs. Rottman and Kelemen call the position advocated by Bering, Barrett, and McCauley, among others, "the strong naturalness thesis." They endorse the naturalness of religion, in that, religious ideas, on their view, are easily assimilated by children, and in that this is not predicated on educational and social factors. But they believe those other accounts are more deterministic than the evidence allows and suggest that an alternative approach, which they call "the development constraints theory," matches more closely what the empirical evidence has revealed so far about the degree in which religion is natural (2012, 9). On this proposal, religious thinking, while emerging from built-in biases, is not an initial stance in children, but something that awaits for a previous

understanding of the world. On this picture, then, there are both nativist and constructivist aspects to the emergence of religious thought in children. Nativist “insofar as it emphasizes some canalization and predicts that religious beliefs and concepts are heavily constrained by initial cognitive biases, reliably unfolding developmental processes, and recurrent environmental inputs” (Ibid, 10). And constructivist insofar as “children are proposed to actively build religious ideas from their experiences in the world” (Ibid). Still, this is a proposal that endorses the naturalness of religion. Children do not embrace religious ideas because of social and cultural factors. Rather, children’s receptivity to religion, while developmentally constrained, is due fundamentally to their creationist and teleological biases.

So, there is broad agreement among CSR researchers that religion is (in some sense) natural. And there is broad agreement that the term “natural” should not preclude the environment, the typical, regular environment where humans grow up here on earth, from being an indispensible component of the development of natural traits. As Boyer put it: "[. . .] if we were all brought up in a sterile and nonmusical environment, we would [not catch colds nor remember tunes]. We would still have the disposition to catch them but no opportunity to do so" (2001, 4). Likewise, developmental and social factors may also constrain and shape the contours of one’s cognitive biases. And as CSR research finally advances through one of its last frontiers (China) to test the naturalness thesis in a challenging and crucial territory, researchers have brought good news for proponents of the naturalness of religion thesis. Hornbeck et al. (2017) report that the thesis “has not merely survived its encounter with the Dragon, it is better for the experience.” But they also acknowledge that the Dragon has reinforced the perception of many that naturalness is not an easy and straightforward concept as one would initially suppose. Is religion natural? “Yes,”they answer, “but comparable to how a gardener can prune, train, water, and nourish a tree to dramatic effect, cultural factors and human creative agency will cultivate our natural propensities to richly diverse forms of expression”(5).

THREE APPROACHES TO THE NATURAL ORIGINS OF RELIGIOUS BELIEFS

In their article “Reformed Epistemology and the Cognitive Science of Religion,” Clark and Barrett (2010) discuss some ways in which the findings of cognitive science of religion and

the explanation offered by reformed epistemologists for the origins of justified religious beliefs have converged. They discuss three main CSR approaches for the origins of religious beliefs and proceed to show how the *sensus divinitatis* postulated by reformed epistemologists can be understood within the framework provided by each of those CSR approaches. In what follows, I present those three approaches. This is important, since in the subsequent chapter we will examine the claim that cognitive science poses a problem for evidentialism (and, in particular, that CSR poses a problem for religious evidentialism). In order to examine this claim, we need clearly stated approaches to CSR that can be used – as Clark and Barrett seem to do – to argue that the findings of CSR are somehow inconsistent with evidentialism. As a result, I will lay out, in the remainder of this chapter, the three approaches used by Clark and Barrett to explore the ways in which CSR and reformed epistemology converge. Since Clark and Barrett wrote that article almost ten years ago, I will complement their discussion with more recent findings in support of those accounts.

1. Attribution Account

1.1. HADD + ToM

We saw above three sets of experiments that show that humans seem to be endowned with an agency detection cognitive faculty (or set of faculties). We saw that newborns are able to discriminate between the biological and the non-biological patterns of motion and attend preferably to biological motion and that that seems to indicate that human beings are already predisposed to detect agency since their first two days of life. We also saw that infants attribute agency to ambiguous stimuli. Three-month-old infants attribute agency to a self-propelled inanimate objects when cues are given that the object is an intentional agent. And when they see an inanimate object moving without the obvious presence of an agent, they expect the presence of a hidden agent.

These and other findings seem to suggest that humans are "eager detectors of agents." In fact, we not uncommongly find ourselves easily identifying objects and entities as agents on scant evidence. A noise or shadow in the middle of night, or a particular formation of a cloud, can be enough to trigger belief that there is a burglar downstairs, or that there is a face in the cloud. For this reason, this mechanism has been termed "hypersensitive." We are

hypersensitive agency detectors. Even the slightlest evidence of the presence of agents are enough to trigger belief that there is or at least there might be another agent around. A natural explanation for the origins of such cognitive mechanism, now called by many "hypersensitive agency detection device" (HADD), is that it provided an evolutionary advantage to our ancestors who possessed them. For, as the saying goes, "better to err on the side of caution."Hypersensitive detectors of agencies had the evolutionary advantage of being less likely to be caught by surprise by predators or enemies.

HADD has become a formidable tool to explain human belief in supernatural agents. Since we are eager – and even hyper – agency detectors, prone to detect agents even when their presence is ambiguous or potentially hidden, it is natural that the hypersensitivity of such a mechanism would be used to explain the human propensity (attested by any brief survey of the history and geography of human cultures) to believe in supernatural agents – God, gods, spirits, ghosts, and so on – when conditions triggering belief are inconsistent with the presence of other known natural agents – such as humans or animals.

But belief in supernatural beings must involve something else. As we saw in our discussion of omniscience and the difficulty children have in understanding the notion of false belief, several other experiments point to the existence of another cognitive mechanism, known as theory of mind (ToM), involved in agency recognition. But while HADD detects the presence of agents, ToM has the function of detecting their mental states, their goals, purposes, intentions. This cognitive mechanism is expected to work together with HAAD. The latter detects agents; the former, their desires, beliefs, and intentions.

1.2. eToM

Bering (2002) has developed an extended formulation of the theory of mind that allows for a broader range of triggers of attributions of mental states about supernatural beings. This broader ToM "serves not to explain or predict behavior but, rather, to allow individuals to attribute meaning to certain classes of autobiographical experiences" (Bering, 2002, 4). While ToM operates in the domain of behavior, eToM operates in the domain of experience. While ToM operates in response to changes in the environment, eToM operates in response to a life event that elicits an experience of meaning, suggesting that such an event is part of

a bigger cosmic scheme. The typical question that emerges in the operatio of ToM is “Why did this happen to me, of all people?” In fact, Bering believes that eToM will elicit intentional attributions by both believers and nonbelievers. The difference is that , while a theist’s answer will likely involve God’s will for her life, the atheists’s answer, on the other hand, will simply involve different elements, such as fate and immanent justice.

Several recent experiments have provided empirical support for eToM. In their article “Everything Happens for a Reason,” Banerjee and Bloom (2015) reported the findings of three experiments aimed at investigating whether children endorse teleological explanations for significant life events. Banerjee and Bloom presented participants from three age clusters (eighty children with ages ranging from five to seven, seventy-two eight-to ten-year-olds, and ninety-one adults) with choices between teleological and nonteleological responses to hypothetical significant life events and explored how helpful children would consider these responses to be for understanding the cause of a significant life event. In the first experiment participants from the three age groups heard descriptions (accompanied by cartoons) of twelve life events, such as “Brianna’s pet puppy ran away from home.” They were then introduced to two fictional characters who offered explanations for the events. Both charecters would initially endorse a nonteleological explanation (“Because Brianna left the front door open”) and at a certain point one of them would offer a teleological explanation (“Because it was meant to teach Brianna that looking after a pet is a big responsibility”) in addition to the first explanation, while the other would deny that there was another, teleological explanation to the event. Other control explanations were also offered and the order and format of the dialogue varied to guarantee the methodological accuracy of the experiment. Banerjee and Bloom reported that both groups of children favored purpose-based explanations, more so than adults, and the first group did so to a greater extent than the second goup, of older children. This suggests that teleological reasoning about life events has roots in childhood but diminishes with age.

In another set of experiments, Banerjee and Bloom (2014) found support for the thesis initially advanced by Bering that even atheists tend to reason teleologically about significant life events. More precisely, they explore in one study the propensity of both believers in God and non-believers to believe in fate, and, in the second study, their propensity to infer

causes to their own life events. Banerjee and Bloom found that even though believers in God tend to hold stronger teleological beliefs than nonbelievers, nonbelievers also perceive design and purpose in life events. Other studies have found additional support for the view that humans are inclined to believe in fate (Norenzayan and Lee, 2010; Tang et al., 2012), and support was also found for the idea that humans seem to naturally embrace ideas about cosmic justice and even karma (Banerjee and Bloom, 2017; Callan et al., 2006; Raman and Winer, 2004).

In summary, the attribution account consists of HADD + ToM + eToM. Clark and Barrett (2010) note that this is, of the three accounts under consideration, the only one that is primarily experiential:

> [The] god-faculty as the Attribution Account suggests [. . .] is activated by specific experiences, and godbeliefs are not pre-existing tacit assumptions waiting to be activated, but are constructed in response to particular environmental stimulation. These experiences could be bumps in the night, faces in the clouds, or striking cases of fortune or misfortune (179).

2. Dispositional Account

The dispositional account is derived primarily from the work of anthropologist Pascal Boyer. According to Boyer, the reason why certain ideas have been so widespread among different cultures and in different periods of history is because they are more easily remembered and transmitted. This is particularly the case of religious ideas. What makes religious ideas more easily remembered and transmitted is that certain kinds of religious concepts are more successful in activating certain mental systems. Religious concepts violate certain expectations about ontological categories (i.e., they are counterintuitive) while preserving other categories, and this makes them particularly attention-demanding and transmissible. To see why certain religious ideas (gods, God, spirits, ghosts, etc) have become so pervasive, consider a candidate for religious idea that did not become common currency among the world religions:

> (a) Some people can see the future but they then forget it immediately (2001, 72)

(a) is counterintuitive – it includes violations of expectations about ontological categories, in this case that minds have representations of events that have not occurred yet. But this idea is not a good candidate for a successful religious concept that can be easily remembered and transmitted. For successful religious concepts must also preserve the relevant default inferences. For instance, a ghost is a person with counterintuitive physical properties. And if you believe you have seen a ghost "your mind creates a whole lot of assumptions of which you are not necessarily conscious. [. . .] You assume that the ghost saw you [. . .] knows [. . .] remember[s] . . .” (2001, 72). The preservation of default inferences is important in that they put the mind in the position to provide information that completes the fragmentary elements initially given, making the idea more memorable. Thus, (a) is not a successful religious concept because, although the idea that people can see the future is counterintuitive – for it violates expectations about ontological categories –, the fact that they forget it immediately precludes its inferential potential from being actualized. So here is Boyer’s essential recipe for supernatural concepts: there must be violation of expectations and there must be inferential potential.

Justin Barrett (2008) argues that the reason why some sorts of beings are more likely to be considered gods than others reflects the fact that they possess a larger number of features that make it easier for humans to conceptualize them as gods. The five features of successful god concepts are: (1) counterintuitiveness, (2) intentional agency, (3) possession of strategic information, (4) ability to act in the human world in detectable ways, and (5) ability to motivate behaviors that reinforce belief (2008, 149). With these five criteria for godhood (so to speak) in hand, he goes on to ask why Santa Claus, in particular, is not considered a god. His answer is that:

> at least part of the reason that the Santa Claus concept is not as widely believed in as god concepts and does not qualify for a prominent role in a religious cult is that Santa does not fully meet the five criteria for being a god. He comes close, but does not quite make the grade. Santa is only sometimes represented as

> counterintuitive. Santa is an intentional agent. His information is marginally strategic at best. He does act in the real world, but only once each year. And Santa only minimally motivates behaviors (mostly during just one month each year) that might reinforce belief. Consequently, Santa Claus does not matter most of the year, he fails to possess rich inferential or explanatory potential, he does not demand attention or speculation, and he is not easily linked to moral or social concerns. In short, he makes for a poor god (2008, 158-9).

So Santa is not a god because the concept of Santa fails to meet all those criteria (or at least to a combined degree needed for goodhood): he is not counterintuitive enough; he does not possess strategic information to a degree that would make him relevant for our survival and reproduction (he supposedly knows which gifts we desire, bringing them to us once a year, but that is of little relevance for our day-to-day life and our survival and reproduction is not seen as depending on it); he acts in the world and in a detectable way (indirectly, via Christimas gifts), but not as frequently as might be needed for godhood; and the Santa concept fails to motivate behavior that reinforces belief. But, while the concept of Santa Claus does not fully satisfy the criteria necessary for a concept to cross the threshold of godhood, it comes closer to it than the concepts of other widespread cultural characters, such as Mickey Mouse and the Tooth Fairy. On the other hand, gods of various religious traditions (*e.g.*, the Christian God, the Hindu Ganesha, and the Maya forest spirits) meet all five criteria suggested by Barrett.

Clark and Barrett (2010, 180) reconstruct Boyer's account by specifying four main cognitive mechanisms that increase the likelihood that certain ideas will emerge, spread, and endure. On this account, we are disposed to believe certain religious ideas when:

(1) They "are easily and readily represented by human cognitive equipment" (2010, 180):

The gods and other supernatural agents people believe in strongly resemble other intentional agents they believe in. Thus when people form the idea of a god or other supernatural agents their ToM is activated and they become disposed to draw inferences

about the agent's mental states and intentions that are similar to those of other intentional agents. To use the ghost example once more, if a child forms the idea of a ghost, he will infer that it has beliefs, can move and see him, can know and remember things. He doesn't need to be taught any of these properties of the ghost. Once he understands that the ghost is an intentional agent, he will infer that the ghost has those properties;

(2) They "are attention-demanding regardless of cultural conditions" (Ibid):

And, as explained above, they are attention-demanding when they violate expectations about ontological categories, *i.e.*, they deviate from our expectations about what there is and how what there is are. The idea that there are agents (spirits and ghosts) who can walk through walls is counterintuitive and attention-demanding. This makes belief in the existence of those agents more easily remembered and transmitted;

(3) They "have rich 'inferential potential' such that they readily generate inferences, explanations, and predictions relevant to many domains of human concern" (Ibid):

As explained above, counterintuitiveness is not enough. If the religious concept is to have mnemonic and transmission advantages, it also has to have rich inferential potential. A person who can see the future has more inferential potential than a person who can see the future but then forgets it immediately. Likewise, ghosts who can walk through walls have greater inferential potential than ghosts who exist for just a second;

(4) They "motivate actions that reinforce belief, *i.e.*, they matter" (Ibid):

When the religious ideas are personally relevant for our lives, in terms of our moral values, security, subsistence, health, etc., they are more likely to be remembered and transmitted. Supernatural being that have properties that may be interesting but that are irrelevant to our lives are less likely to become part of our religious repertoire. For instance, gods or spirits who can see the future but only with respect to trivial matters that won't affect our lives possess counterintuitive properties and provide inferential potential, but have little or no relevance for our moral or practical lives and thus are less likely to become part of our religious views than gods or spirits who can see the future, including our own future, and can, thus, be sought to assist us with our problems.

Thus, on the dispositional account, certain religious ideas, in particular idea about supernatural agents such as God or gods, meet certain conditions that make those ideas relatively easy to memorize and be transmitted. This is, according to proponents of the dispositional account, the primary reason why certain religious ideas are so prevalent and pervasive. God concepts are easily and readily represented by the human cognitive equipment, are attention-demanding, have rich inferential potential, and are important enough to motivate action and reinforce belief. Unlike what the attribution account tells us, here experiences are not needed to trigger belief. This model is epidemiological: the human mind is particularly susceptible to being infected by certain religious ideas. Once we understand those ideas we are inclined to "embrace them, act upon them, and pass them along" (Clark and Barrett, 2010, 182). The inputs here are, therefore, not experiential, according to Clark and Barrett. The outputs of these dispositions when it comes to god beliefs, on the other hand, are vague: "the god will not be identical with a human (or it would not be attention-demanding and have richer inferential potential), but it could have any number of properties such as being invisible, superknowing, superperceiving, superpowerful" (Ibid.).

3. Preparadness Account

Jean Piaget (1972) advanced the thesis that children think that everything that exists in the natural world is made by people for a purpose. Many of the children he interviewed attributed the existence of natural entities to humans. He called this tendency in children "artificialism" and concluded that the reason why children are artificialists is because they lack the conceptual capacity to conceive physical causes and thus cannot distinguish between natural entities and artifacts. In addition, Piaget noted that children had the tendency to attribute omniscience and omnipotence to humans and that they often represented God with human-like properties. He also noted that children often faced a "crisis" when they realized that there were things that their parents did not know, but often retained belief in a human-like deity that knew everything. As Barrett and Ritchert summarize Piaget's views about the way children understand God: "[for] Piaget God is a parent who fulfills intellectual needs to account for the structure of the world" (2003, 303). Therefore, after children realize that their parents are not omniscient and omnipotent, they

retain belief that there is an all-powerful and all-knowing human-like being in the sky. Children, according to Piaget, lack the conceptual apparatus to conceive God more abstractly. This view became known as “anthropomorphism.” The anthropomorphism thesis has been embraced by many psychologists and cognitive scientists and has led many of them to conclude that children lack religiosity.

These views have been challenged by subsequent research. Both artificialism and anthropomorphism have been shown to rest on problematic assumptions and on inadequate research methods. Subsequent research has shown that children can in fact reason in terms of physical causes and recognize that humans, not natural causes, make artifacts (Gelman and Kremer, 1991; Keil, 1989; and Rottman and Kelemen, 2012). Children can also represent God abstractly, in non-human-like forms. Barrett and Ritchert (2003) and Barrett (2012) summarize a series of studies that cast serious doubts on the anthropomorphism hypothesis (see also, among others, Kelemen, 2004, for criticisms of anthropomorphism). In particular, children seem to reason about God’s attributes of creative power, knowledge, and immortality in ways that are inconsistent with what we should expect if anthropomorphism were true.

The view that the attribution of creative power to God by children is derivative from their attribution of such power to humans has been challenged by studies such as those conducted by Petrovich (1991, 1997). In one of such studies, Petrovich presented children with a pair of photos of animals, plants, other natural kinds, such as snow, toy animals and plants, as well as artifacts, such as chairs. The children were then asked whether the content of the photographs was “something that can be made by people or something people can’t make” (1991, 10). The children’s answers were remarkably accurate when the pairs displayed clear cases of natural kinds and artifacts, as opposed to toy animals and plants. They had little difficulty distinguishing objects that could be made by people from those that could not. The conclusion that has been drawn from this and other studies is that children can in fact discriminate natural kinds from artifacts. In addition, studies have shown that children strongly favor creationist views over evolutionary explanations for the origins of the natural life-forms, even when there is control for the religiosity of the parents (Evans, 2001).

Children, as illustrated by the "false beliefs tasks" mentioned above, have difficulty understanding that beliefs can vary from person to person and that other persons (including imaginary friends, animals, plants, and dolls) can have false beliefs until they are about five-years old. But even as older children acquire the capacity to understand that other persons can have false beliefs, they continue to think that God knows what is truly inside the box. Barrett and Richert and others draw the conclusion from this and other findings that the idea that "beliefs are infallible" (2003, 304) is the default assumption for children, something that does not disappear naturally, but, rather, needs to be unlearned.

We have also seen that several authors believe that belief in immortality – or at least that we are embodied souls that will continue to live after death – is a default assumption of children. In addition to the conclusion that children view themselves as immortal, studies show that they also see other agents, including God, as having the property of immortality. Such assumption, like the omniscience assumption, also is an assumption that needs to be unlearned.

And, as also mentioned above, children also assume that God is super-powerful – and perhaps even all-powerful. According to Piaget, this is so because children attribute super-strenght to humans and, since God is just another sort of human, but located in the sky, children also attributes super-power to God. In his interviews with children, Piaget noticed that children attributed super-power to their parents in addition to attributing the creation of the sun, mountains, and so on – which requires super-power, no doubt – to humans and to God. And as we have seen in the case of omniscience, eventually children learn that humans can have false beliefs while retaining the belief that God's beliefs are "infallible." Likewise, suggests Barrett (2012, 74-77), they learn that humans have limited power, but seem to retain belief that God's power is not affected by the same constraints.

Recent research, however, seems to vindicate at least one aspect of Piaget's account: children do reason teleologically about natural entities. When asked why rocks are pointy, children often give teleological answers such as "so that animals will not sit on them" (Kelemen, 1999). Or, when asked why a certain tiger exist, children frequently give answers that manifest their teleological bias, such as "so that it can be seen in the zoo." The empirical evidence about children's propensity to reason teleologically is robust. This has led, as we

saw above, to further research to explore the hypotheses that this teleological propensity remains in adults – even scientifically educated ones – and that it is universal (Kelemen and Rosset, 2009). In their article suggestively entitled "Professional Physical Scientists Display Tenacious Teleological Tendencies: Purpose-Based Reasoning as a Cognitive Default," Kelemen, Rottman, and Seston (2013) present the results of a study with physical scientists from top-ranked American universities that show that they tend to reason teleologically when their information-processing resources are limited. It has also recently been shown that the thesis holds with respect to non-western subjects socialized in cultures (as in the case of China) that do not have been under the cultural influence of Abrahamic theism, and in fact, have been strongly influenced by an atheist worldview and that has only a small percentage of the population claiming to believe in God (Rottman et al., 2016, Hornbeck et al., 2017, Järnefelta et al., 2018).

The account that has emerged from these and other findings has been termed by Barrett and Ritchert the "preparedness thesis." In contrast to the anthropomorphic approach, which claims that children lack religiosity and that God concepts arise from a propensity to attribute to humans what should be attributed to physical causation, the preparadeness approach tells us that religiosity and God concepts arise from a cognitive propensity to see things in the world as existing for a purpose. This propensity is so strong and widespread that its robust confirmation by empirical findings has led to a debate on whether children are "intuitive theists" (Kelemen, 2004).

It is important to note that Kelemen herself, as we saw above, does not view the human capacity for religious belief as being something that is "prepared early on" (2012, 2) and, thus, distances herself from the label of "preparadness." Rather, she views religious belief as being "something that develops over the course of the first several years of life," in combination with "biases that initially prepare young children for thinking about the natural world, recurrent experiences and inputs, constructivist tendencies, explanatory motivations, and other important components of typical development, children predictably begin to latch onto religious ideas by middle childhood" (Ibid).

Clark and Barrett (2010) note that the preparadness approach is similar to the dispositional approach in that the cognitive mechanisms involved in the formation of religious ideas or

beliefs in these approaches are not triggered by experiences (as it is the case of the attributional approach). Rather, it is part of our cognitive development to come to "see" the world and its constitutive entities as intentionally, purposefully designed. As Clark and Barrett put it,

> Children automatically see the natural world as having intelligent, intentional design and purpose and are prone to see the creator as superknowing, superperceiving, and immortal. The only 'triggering' necessary is for the particular name of the Creator to be specified (2010, 184).

In Barrett's terminology, we are "prepared," by our evolutionary development, to come to see the world as children as the product of (super-powerful and super-knowing) agency. And this tendency largely remains in adults, even in highly educated ones.

REFORMING REFORMED EPISTEMOLOGY?

In this chapter, we will bring together some of the central questions discussed in the previous chapters on the nature of evidence, on evidentialism, and on the main findings of cognitive science of religion, in order to evaluate the first claim that concerns us in this essay: cognitive science indicates that evidentialism about epistemic justification is false and CSR, specifically, indicates that religious evidentialism is false.

In order to properly examine this claim, we need to explore whether what cognitive scientists of religion are saying about the naturalness of religious beliefs is incompatible with evidentialism or at least with the idea that the subject that forms justified religious beliefs on the basis of the mechanisms described by CSR can plausibly be said to be responding to evidence in some way or another. We do this in section two below, by returning to the discussion of the previous chapter about the three CSR accounts of the naturalness of religious beliefs (the attribution account, the dispositional account, and the preparedness account). As we will see, one can plausibly suppose that subjects can form (justified) beliefs under the mechanisms described by these accounts and can do so by forming such beliefs in response to evidence and are perhaps justified in their responses. This, however, may not be enough to show that claim one is false, for there may be objections that are unrelated to those three accounts to the effect that CS or CSR show that evidentialism is false. And, in fact, there are such objections: McCauley (2011), Greco (2006), and Barrett (2004, 2009) have presented objections that are unrelated to the main mechanisms especified by the three CSR accounts, relying on more general findings from cognitive science. If successful, these objections strongly indicate that cognitive science shows that evidentialism about epistemic justification is false. Hence, in section three we present their objections and reconstruct them in the form of valid arguments. As we will see, each of the arguments relies on problematic premises and fails to show that the findings of cognitive science about how the human mind operates are incompatible with evidentialism. But before we do this, we have to come to a proper understanding of what reformed epistemology and the *sensus divinitatis* are and explore whether they can be explained in terms of an evidentialist account of justification. For, as Clark and Barrett (2010) noted, reformed epistemology and CSR have remarkably converged. Their understanding of reformed epistemology, however, is

clearly along the lines defended by Alvin Plantinga, *i.e.*, an externalist and proper functionalist account. We need, however, to explore whether an evidentialist account of reformed epistemology can be properly formulated so that we can ascertain whether, at least in principle, (evidentialist) reformed epistemology and CSR can also converge. We will do so by, first, examing Plantinga's account of reformed epistemology and of the *sensus divinitatis*. We will then see two attempts that have been made (by Tucker, 2011, and by McAllister and Dougherty, 2018) to develop accounts of reformed epistemology that are consistent with internalism and evidentialism. As we will see, there is no obstacle for the development of an account of reformed epistemology in terms of an evidentialist approach to justification and, as a result, whether CSR and evidentialist reformed epistemology converge is just a matter of the findings of CS and CSR being compatible with what evidentialists can plausibly say about evidence possession and evidential support.

Finally, the chapter ends with a brief exploration of the prospects for the development of improved accounts of reformed epistemology and of the *sensus divinitatis* in light of our examination of plausible alternative evidentialist accounts that can accommodate the findings from the three CSR accounts examined in section two.

1. Reformed Epistemology

1.1. Proper Functionalism and the Sensus Divinitatis: Plantinga's Model

What is now kown as reformed epistemology emerged in the second half of the twentieth century as a response to classical evidentialism.[38] Classical evidentialism held that a belief is

[38] The main architects of reformed epistemology were Alvin Plantinga, Nicholas Wolterstorff, and William Alston. The version of religious epistemology defended by these philosophers was named "reformed" because Plantinga and Wolterstorff initially made explicit reference to the Christian reformer John Calvin as one of the forefathers of the main tenets of the epistemology they were defending. However, while Calvin and other reformed thinkers do seem to have endorsed the main tenets of what would eventually be called "reformed epistemology," those ideas have by no means been exclusively embraced by reformed thinkers. Aquinas, for one, held similar views, as evidenced by the fact that Plantinga himself calls his model of warranted theistic belief the "Aquinas-Calvin model." And William Alston favored the expression "episcopalian epistemology" to refer to the epistemological theory he helped to develop along with Plantinga and Wolterstorff. In fact, it seems that what is now known as reformed epistemology has been the predominant view among Christian thinkers of all the main Christian denominations about the rational grounds of theistic and Christian beliefs since the advent of Christianity, and the view embraced, as we will see below, by nearly all (if not all) contemporary theist epistemologists.

justified only if it is properly basic or results from an argument whose premisses are properly basic. Properly basic beliefs are, according to the classical evidentialist picture, beliefs about propositions that are certain (incorrigible) or self-evident. Incorrigible beliefs are those about one's own mental life: that the flower, as it appears to me, is yellow, or that I'm in pain. If a subject is having the experience that the flower appears (as opposed to "is") yellow or that she is in pain, then belief in those propositions are certain or incorrigible to that subect. And a belief is self-evident if it is so obvious that one cannot even understand it without seeing that it is true. The belief that 2+2=4 or that there is no man taller than tallest man in the world are examples of self-evident propositions.

On this picture, it has been claimed, belief in the proposition "God exists" is not justified in the basic way, for it is not incorrigible, nor self-evident. Therefore, theistic belief can only possibly be justified on the basis of propositional evidence, i.e., on the basis of arguments whose conclusions follow from properly basic premisses. Reformed epistemologists' main goal is to show that theistic beliefs can be properly basic and, therefore, that arguments from natural theology are not needed for the rational justification of theistic beliefs. The reformed epistemologist's case for the proper basicality of theistic beliefs has involved two claims. First, they seek to show that the claim that only beliefs that are incorrigible or self-evident are justified is self-refuting. They do so by pointing out that belief in classical evidentialism itself cannot be incorrigible or self-evident. Therefore, by its own criteria, belief in classical evidentialism cannot be justified. Second, they provide what they take to be a plausible story about how theistic beliefs can the held in a properly basic way. The most prominent such account comes in the form of Plantinga's Aquinas/Calvin model, or A/C, for short. According to this model – which Plantinga develops from what both Aquinas and Calvin said about the ways in which humans come to know God – God designed us in such a way that we have a natural inclination to believe in and know him. Aquinas writes that "To know in a general and confused way that God exists is implanted in us by nature."[39] And Calvin seems to concur:

[39] Summa Theologiae I, q. 2, a.1, ad 1.

> There is within the human mind, and indeed by natural instinct, an awareness of divinity. This we take to be beyond controversy. To prevent anyone from taking refuge in the pretense of ignorance, God himself has implanted in all men a certain understanding of his divine majesty. . . . Since, therefore, men one and all perceive that there is a God and that he is their maker, they are condemned by their own testimony because they have failed to honor him and to consecrate their lives to his will . . . there is, as the eminent pagan says, no nation so barbarous, no people so savage, that they have not a deep seated conviction that there is a God. . . . Therefore, since from the beginning of the world there has been no region, no city, in short, no household, that could do without religion, there lies in this a tacit confession of a sense of deity inscribed in the hearts of all.[40,]

Thus, on Calvin's account, this knowledge of God is the result of a natural inclination to have a sense of deity. Following Calvin, Plantinga called this sense of deity "the *sensus divinitatis*." Unlike Calvin, however, Plantinga defended the view that the *sensus divinitatis* is more than a sense or perhaps inclination to believe, but a cognitive faculty that gives us a more well-defined object of belief – the theistic God or something close to it. Thus, on Plantinga's model, just as we are endowed with cognitive faculties such as perception, consciousness, reasoning, memory, and so on, we possess a God-faculty designed to give us knowledge of God.[41] Unlike Calvin again, who conceived the *sensus divinitatis* as having as inputs our dispositions to believe in a divine reality of some sort, requiring no special experiences, and

[40] Institutes I, iii, 1, p. 44.

[41] This account of the naturalness of belief in God found in these authors is taken to have Scriptural support, primarily from Romans 1:18-20, where Paul says that: "For the wrath of God is revealed from heaven against all ungodliness and unrighteousness of men who suppress the truth in unrighteousness, because that which is known about God is evident within them; for God made it evident to them. For since the creation of the world His invisible attributes, His eternal power and divine nature, have been clearly seen, being understood through what has been made, so that they are without excuse." While many have taken this passage to endorse natural theology, Plantinga, following Calvin, believes this passage is more accurately understood as supporting our natural inclination to believe in God under certain circumstances, such as when we find ourselves in the presence of the grandeur and splendor of nature.

as output a vague sense of a not well-specified divinity, Plantinga sees its inputs as consisting of special experiences, and its output as consisting of more well-defined theistic beliefs (see Clark and Barrett, 2010, for discussion of this distinction).

Plantinga's account of the *sensus divinitatis* is informed by his epistemomological externism and his proper functionalism about warrant[42]: the *sensus divinitatis* can deliver warranted theistic and Christian beliefs even if the subject is not aware of how her belief is connected to her evidence, or even when she is not responding to her evidence, or perhaps even when she doesn't have any evidence available for the belief that is being formed. What matters on his account is that the belief meets the conditions for warrant. And, according to proper functionalism, in order to have warrant, a belief has to be produced by cognitive faculties that are functioning properly (that is, free from malfunction or dysfunction). Proper function involves the notion of design plan (by evolution or God, or both): the cognitive faculties work properly if they are working as they should work. Like other of our human organs, there are ways in which our cognitive faculties are supposed to work. A properly functioning human heart, for instance, should beat about 50-80 times a minute when we are at rest. Moreover, the design must be a good one in the sense that the purposes of the design plan will be achieved. In the case of the human heart, its purpose is to pump the blood. Likewise, cognitive faculties can function properly or malfunction. The main component of a proper functionalist account of warrant is, therefore, that the cognitive faculties are functioning properly, subject to no significant dysfunction.

But there is more. Our human organs are designed to work in certain environments. We – to use some of Plantinga's examples – can't breathe under water; our muscles atrophy in zero gravity; we can't get enough oxygen at the top of Mt. Everest.[43] Likewise, cognitive faculties will function properly only in environments for which they were designed. They wouldn't work well, for instance, to use another of Plantinga's examples, in an environment in which "a certain radiation impedes the function of memory."[44]

[42] Warrant is that, whatever precisely it is, which makes the difference between knowledge and true belief.

[43] Plantinga, Alvin (2015) *Knowledge and Christian Belief*, position 686 of Kindle.

[44] Ibid.

In addition to cognitive faculties working according to a good design plan and in an environment for which they were designed to work, the purpose of cognitive faculties functioning properly is to produce true beliefs. Such cognitive faculties must, therefore, be aimed at producing true beliefs. And since the design plan is a good one, there has to be a high probability that beliefs produced according to the plan will be true.

Thus, as Plantinga summarizes his account of warrant,

> A belief has warrant for a person *S* only if that belief is produced in *S* by cognitive faculties functioning properly (subject to no dysfunction) in a cognitive environment that is appropriate for *S*'s kind of cognitive faculties, according to a design plan that is successfully aimed at truth.[45]

As a result, on Plantinga's views about warrant, the *sensus divinitatis* can deliver warranted theistic and Christian beliefs if God designed us to form theistic and Christian beliefs via such a faculty, and this faculty is functioning properly, was triggered at or by the right sort of environment, and the subject's belief is aimed at truth. Thus, like our other human cognitive faculties, the *sensus divinitatis*, when working properly, will produce warranted theistic beliefs, *i.e.*, beliefs in God and about God that, if true, will amount to knowledge. Beliefs cannot be warranted when they are not aimed at truth, when they are motivated by, for instance, psychological comfort or desire for self-esteem. Likewise, when a *sensus divinitatis* that is not working properly, due to the effects of sin (such as pride, fear, envy, lust, and all the panoply of human deviations from what is good), our capacity to know God and his attributes becomes inhibited.

1.2. Seemings and the Sensus Divinitatis: Tucker's Model

We have seen that Plantinga's reformed epistemology is grounded in his externalist and proper functionalist views about warrant. The central tenets of reformed epistemology, however, can arguably be incorporated into alternative approaches to epistemic justification and warrant. In fact, reformed epistemology, as defined by its main proponents, seems to be

[45] Ibid.

neutral with respect to debates such as internalism vs. externalism and evidentialism vs. the views that are inconsistent with the evidentialist thesis. Andrew Moon, for instance, defines reformed epistemology as "the thesis that religious belief can be rational without argument" (2016, 879). And Michael Bergmann, a proper functionalist and reformed epistemologists himself, defines reformed epistemology as "the view that belief in God can be rational even if it is not inferred from any other beliefs" (2014, 613). Thus, on this, predominant understanding of what reformed epistemology is, there is nothing in it that would in principle preclude an internalist or evidentialist about religious epistemology from claiming the label for herself. But perhaps these understandings of reformed epistemology are missing an important component of what some of the main proponents of reformed epistemology have been advocating. For, as mentioned before, it is crucial in the formulation of the religious epistemology of Calvin, Plantinga, and others, that justified or warranted theistic beliefs are formed not only non-inferentially or without the need for arguments of natural theology, but by a specific sense or mechanism or faculty, the *sensus divinitatis*, which *naturally* delivers rational religious or theistic beliefs. Thus, perhaps the following definition is closer to the spirit of Calvin's and Plantinga's understanding of what is involved in the formation of rational and warranted religious or theistic beliefs:

> *Reformed Epistemology* (RE*):* the thesis that religious (or perhaps, more specifically, theistic) beliefs can be rational or warranted without argument and formed naturally by means of a sense, or mechanism, or faculty that is triggered when we find ourselves in certain circumstances.

I find this definition more in the spirit of Calvin's and Plantinga's proposals and I believe it avoids the strange result delivered by the definitions put forward by Moon and Bergmann of characterizing all (or nearly all) theist religious epistemologists of the past and of today as reformed epistemologists (many of whom have no association with the reformed tradition in theology).[46] Yet, my goal in laying out this alternative definition is not to defend it here, but

[46] On Moon's and Bergmann's definitions, even Paul Moser (2008, 2010), whose religious epistemology is predicated on the idea that theistic beliefs are justified and amount to knowledge when we have direct acquaintance with the evidence (God himself), and, Richard Swinburne (2010, 2018), the paradigmatic natural theologian of our time, would be considered reformed epistemologists. Moser's religious epistemology, however, seems to be incompatible with the thesis of the naturalness of religious beliefs advocated by

just to point out that it seems reasonable to request from any adequate alternative (non-externalist, non-proper functionalist) account of reformed epistemology that it be capable of providing a convincing account of the *sensus divinitatis* and of the naturalness of religious beliefs. With respect to the former, we find in the literature an important attempt to develop an account of the *sensus divinitatis* along evidentialist lines – that of Chris Tucker (2010).

Tucker's account of how religious beliefs can be non-inferentially justified but at the same time justified on the basis of evidence relies on phenomenal conservatism, which he defines as:

> *Phenomenal Conservatism (PC):* if it seems to S that P, then S has evidence which supports P (2010, 4).

Thus formulated, PC is a thesis about evidence possession: it tells us that seemings provide us with evidence. If it seems to me that Mario is in good health, that provides me with at least some evidence that Mario is in good health. Or, to use a non-perceptual example, if it seems to me that I ate one egg for breakfast today, then I have at least some evidence that I ate one egg for breakfast today. A seeming, in the influential characterization of William Tollhurst, "[has] the feel of truth, the feel of a state whose content reveals how things really are" (1998, 298-9). There are different proposals on how to best conceive seemings. Some authors take seemings to be beliefs. Others have suggested that seemings are inclinations or dispositions to believe. But the view that seems to attract the endorsment of the majority of PC proponents today is that seemings are some sort of experience with propositional content or a *sui generis* propositional attitude (see Tucker, 2013, Huemer, 2013a, and Moretti, 2015, for discussion of these views).

And on this picture, the epistemic support a belief receives from the evidence – the seemings – is directly proportional to the strenght of the seeming: the stronger the feel of

---

cognitive scientists of religion (something emphasized by Moser in personal communication) and apparently advocated by Calvin and Plantinga. Swinburne, while acknowledging that his understanding of evidence makes his and the predominant understandings of evidentialism compatible with reformed epistemology's basic tenet that religious beliefs can be justified without argument (2010) (since his version of epistemic conservatism allows for basic beliefs, inclinations to believe, and experiences, to count as evidence), has been silent on whether his views are consistent with the findings of cognitive science of religion and the naturalness thesis.

truth, the stronger one is justified in believing the proposition in question, provided there are no defeaters for the belief. This picture of evidential support is in marked contrast with that advanced by proponents of proper functionalism. On their view, evidential support is a matter of what the proper functioning of one's cogitive faculties would dictate one should believe in a given circumstance. Proper functionalists reject what Bergmann calls necessity

> *Necessity (N)*: the fittingness of doxastic response B to evidence E is an essential property of that response to that evidence (2007, 112).

This means that the fittingness of a doxastic response to the evidence doesn't require the satisfaction of other conditions. And proper functionalists believe, however, that this fittingness is dependent on the design plan of the creatures. Proper functionalists believe that if evolution or God (or both) had designed humans differently we might end up with different fitting responses to a certain body of evidence. Bergmann illustrates this point with an example involving alien cognizers "who form the belief that there is water nearby via a belief-forming process that bypasses their other mental states" (64). Bergmann asks us to suppose that the formation of such belief by the aliens is not only reliable but in accordance with what is proper function for them to believe in those circumstances given the way they were designed. Bergmann claims that this belief, formed in the way it was by the aliens, is intuitively justified, even though *N* was violated and there was no input in the form of other mental states that led to the formation of the belief. Proper functionalists argue, then, that there is no need for the causal intermediation of other mental states for beliefs to be justified or warranted. All that is necessary is the fulfilment of the proper function requirements for justification and warrant, and those requirements are external to the mental life of the subject.

On Plantinga's model of the *sensus divinitatis*, a non-inferential belief in God or about God may likewise be triggered by the sensations without the intermediation of other mental states. Belief in God or about God may be triggered by one's sensation caused by the visual imagery of a sunset (figure 1), without the need for any other intermediating mental state, if that is what a properly functioning design plan. The problem with this formulation of the *sensus divinitatis* that rejects N and the need for the causal intermediation of other mental states is that it strikes many as arbitrary. A sunset in itself doesn't seem to provide epistemic

support for the proposition “God exists” or “God loves me.” After all, if the imagery hás any propositional content, it is “this is a beautiful sunset,” not “God loves me.” And if a sunset in itself can provide suport for the belief that “God loves me,” why not any other visual imagery? This picture of epistemic support in terms of proper function may seem, therefore, counterintuitive to many.

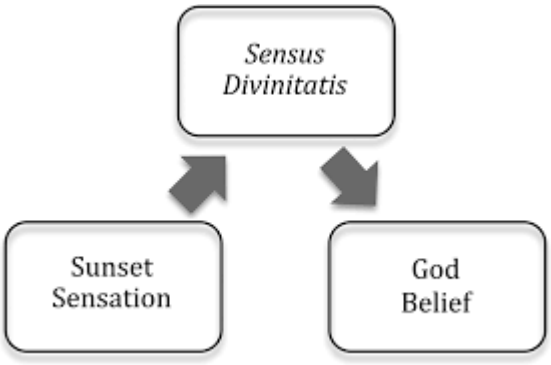


Figure 1 – Plantinga’s Model[47]

In order to avoid these problems, while allowing for the awe and wonder of the obseervation of a beautiful sunset to produce justified and warranted non-inferential theistic beliefs, Tucker constructs a model of the *sensus divinitatis* in terms of phenomenal conservatism. In his reconstruction of the *sensus divinitatis*, it is not the sensation caused by the visual imagery of the sunset that triggers a non-inferential theistic belief, but the seeming produced by the vision of the sunset (figure 2). Rather than having as output a belief, as in the case of Plantinga’s model, here the output is an experience that then provides the grounds for belief formation. Given that seemings are evidence, non-inferential theistic beliefs are thus formed, on this model, in response to evidence. And, all else remaining equal, the stronger the seeming with the propositional content “God loves me” that triggered by the imagery of the sunset, the stronger one is justified in believing in God’s love for her. Tucker’s model is therefore consistent with evidentialism: the sunset can in fact provide evidence for propositions such as “God exists” or “God loves me.” It not only avoids

[47] The figures of the models of *sensus divinitatis* in this chapter are from McAllister and Dougherty (2018).

the problems faced by the proper functionalist acount of the *sensus divinitatis*, but it also makes the acquisition of evidence for religious beliefs something very easy.

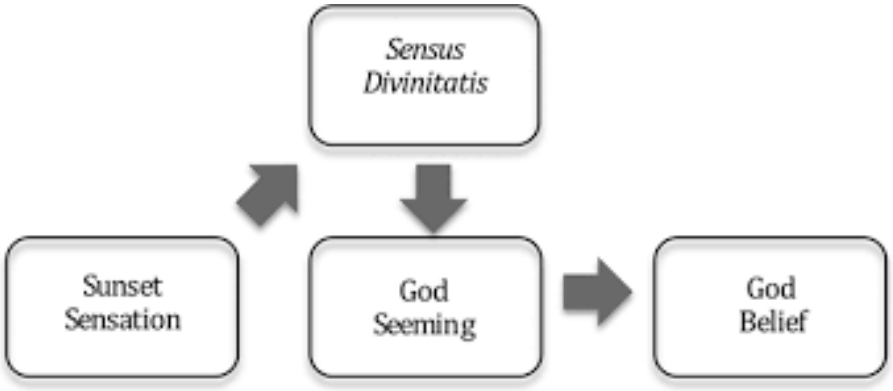


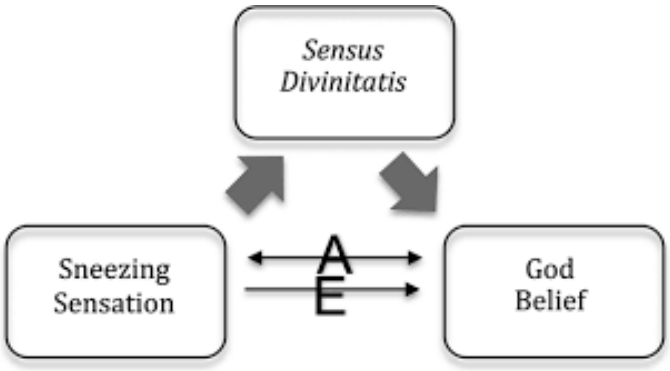
Figure 2 – Tucker's Model

1.3. The Tacit Perception of Evidential Support and the Sensus Divinitatis: McAllister and Dougherty's Model

McAllister and Dougherty (2018) (henceforth M&D) share Tucker's concern for the alleged shortcomings of the model of the *sensus divinitatis* construed in terms of proper function and they see Tucker's attempt to lay the foundations of the *sensus divinitatis* on the grounds of phenomenal conservatism as an important improvement vis-à-vis the previous model. However, they see Tucker's model as still involving reproducing the original arbitrariness in the conversion of the sunset sensation to the respective mental state by the *sensus divinitatis*, with the difference that while the previous model tells us that the *sensus divinitatis* converts a sensation into a belief, Tucker's model tells us that the *sensus divinitatis* converts a sensation into a seeming, which then triggers a belief.

In order to make the arbitrariness that is allegedly present in both Plantinga's and Tucker's models explicit, they replace the sunset by a sneezing sensation in the conversion schema (figure 3, with <-A-> designating arbitrariness). On Plantinga's model, the design plan "programed" the *sensus divinitatis* to respond to a visual imagery of a beautiful sunset with the belief that "God loves me." Suppose now that the design plan likewise programed the

sensus divinitatis to yield the same belief when I sneeze. Since, on proper functionalism, *necessity* is false, a sneeze, like a sunset, could provide evidence for a propostition like "God loves me." Recall that, on proper functionalism, the belief of alien cognizers "that there is water nearby" formed "via a belief-forming process that bypasses their other mental states" can be justified if the belief is reliable and formed in accordance with what is proper function for them given their design plan (Begmann, 2007, 64). On the proper functionalist view of evidential support, a mental state is evidence for a certain proposition if one's belief that *p* is based on that mental state and formed in a way that satisfies the proper function conditions for warrant (Plantinga, 1993, 165 and 168, and Bergmann, 130-31)[48] (In figure 3, -E-> designates the relation betwen a mental state and a belief; with the relation holding if the mental states provides evidence for the belief). So, according to Plantinga's model, arbitrary experiences, such as a sneezing sensation, which are totally unrelated to the existence of God, can provide evidence for theistic propositions and trigger justified and warranted theistic beliefs.



[48] Here is an excerpt from what Plantinga says about this: "What makes it the case, therefore, that B is evidence for A (what makes A epistemically probable with respect to B) is the effect JB has on the degree of belief enjoyed by A in a sound understanding—a sound human understanding, that is; things might go quite differently for Alpha Centaurians or angels" (165; see p. 168 for his proper function account of evidential support).

And Bergmann (2007, 130-1) writes that "the fittingness of a doxastic response to evidence is contingent upon the proper function of the cognitive faculties of the person in question. And this, in turn, suggests that the evidentialist claim E_F [S's belief B is justified iff B is a fitting doxastic response to S's evidence] could be illuminatingly improved if it were changed to say something like the following (where a *PF-induced* doxastic response is one produced by the proper functioning of the subject's cognitive faculties): E_{PF}: S's belief B is justified iff B is a *PF-induced* doxastic response to S's evidence [. . .] epistemic fittingness should be understood at least partly in terms of proper or healthy cognitive functioning."

Figure 3 – Plantinga's Model 2

Tucker's model (figure 4) has the advantage of yielding the result that it is not the sneezing sensation that provides evidence for theistic propositions, but the seeming with a theistic propositional content. Nonetheless, the conversion of a sneezing sensation into a seeming with theistic propositional content is no less arbitrary than the conversion of a sneezig sensation into a theistic belief (figure 4). To be sure, once the subject has a seeming, she has evidence, and then her belief can properly be said to be based on evidence. But this doesn't remove the arbitrariness of the first stage of this causal chain: that the *sensus divinitatis* would turn a sneezing sensation into a seeming is no less arbitrary (lacking in evidential connection) than its turning the same sensation into a belief.

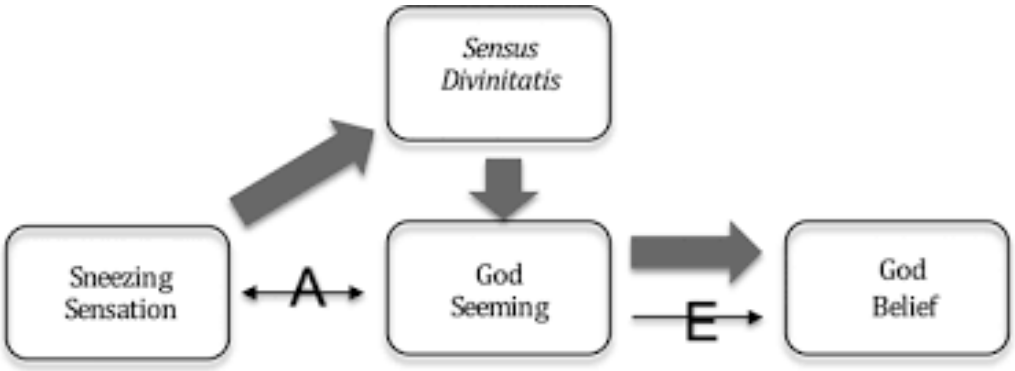


Figure 4 – Tucker's Model 2

So, while there is progress from Plantinga's model to Tucker's, in that seemings can properly be said to be evidence and might justify a theistic belief, both models face the problem of arbitariness. In order to circumvent this problem, M&D defend an alternative model, the Reductive Model, that has the advantage, according to them, of eliminating the counterintuitive way through which the *sensus divinitatis* is said to turn sensations into beliefs or seemings. Their model is built around the notion of tacit perception of support relations between propositions. This is the unconscious ability to perceive in an immediate

(non-inferential) way that one proposition supports another. They illustrate this ability with several examples. In one of them, a husband enters his home, sees his wife's keys on the table and shouts: "Honey, I'm home." M&D's claim is that no conscious reasoning took effect in this scene. The husband immediately took the presence of his wife's keys on the table to be evidence that she was home. Or suppose, alternatively, that the husband entered his home and didn't see his wife's keyes on the table, which was evidence for him that she wasn't home. But he immediately sensed that something was wrong, though he couldn't say what the problem was. He starts to sweat and his heart starts beating faster. He is tacitly detecting elements in the environment that are out of the ordinary. He is, in fact, tacitly processing several pieces of information (the slightly different temperature and the light and wind coming subtly from one of the rooms, for instance) that indicate that somebody broke into the house.

M&D have a lot more to say about tacit perception. It suffices to note for our purposes here, however, that when these tacit perceptions become conscious, they are naturally manifested via seemings (figure 5). "This," M&D suggest, "leads to a natural account of seemings according to which they result from tacit graspings of support relations" (2018, 10). And this provides M&D with the element they needed for an alternative understanding of the *sensus divinitatis* that can potentially avoid the problem of arbitrariness found in the previous ones. For, on this new account, the seeming (unlike the belief of Plantinga's account, and the seeming of Tucker's account) is the direct result of the perception of the evidential connections between propositions. And it is the sunset seeming, not the sunset sensation, as in the case of the two previous models, that triggers the *sensus divinitatis*. Moreover, on the M&D model, the sunset seeming is likely to trigger the *sensus divinitatis* in combination with the subject's background information. Thus the sunset will lead the subject to believe that God loves her, not simply because of the beauty of the sunset, but because the awe and majesty expressed by the beauty of the sunset will fit with several other patters of information the subject possesses. The recognition of the evidential connection between all these bits of information is processed on a tacit, unconscious level, and the subject may never be able to fully grasp on a conscious level all these intricate relations. From the subject's perspective, she may simply find herself with the seeming that

God loves her when she finds herself engulfed by the awe and wonder occasioned by the sunset.[49]

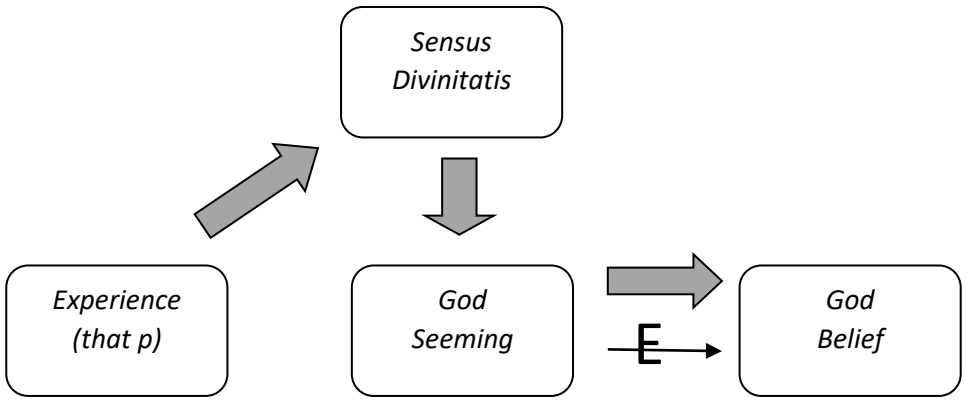


Figure 5 – McAllister and Dougherty's Reductive Model

As a result, the *sensus divinitatis* of this account is not a black box that turns sensations that don't have theistic propositional contents into theistic beliefs or seemings. In fact, the *sensus divinitatis* now is not even a faculty or single mechanism whose function is to produce theisitc beliefs. Rather, as M&D describe it:

> We have a *sensus divinitatis* simply because we have a tendency to draw connections between the content of our experiences and propositions implying the existence of God. Thus, on our model, the *sensus divinitatis* is nothing other than a sub-function of our standard rational faculties (2018, 12).

Thus, the M&D model is reductive precisely because it *reduces* the *sensus divinitatis* to our ordinary faculties. This is an important consideration in favor of their model in that CSR

[49] One worry is that what we have here is a case of unconscious inference, as opposed to the non-inferential belief that motivates reformed epistemology. In response to this M&D say that: "These automatic, sub-personal calculations are not things that we, properly speaking, do. They are things that happen in us. The sense of 'non-inferential justification' which foundationalists have in mind is of justification that does not result from any inference we make. Thus, the existence of an unconscious inference does not endanger our success in modelling non-inferentially justified theistic belief" (13).

seems to indicate that we don't have a unique faculty devoted to the formation of religious beliefs (as Plantinga defended). Rather,

> the widespread view of CSR scholars is that the faculties that incline humans toward religious beliefs are part of the general conceptual toolkit for negotiating life as a human and not some special religion-specific faculty or "god spot" in the brain (Barrett and Church, 2013, 312-313).

Another important result yielded by the M&D model that aligns it well with (at least some of) the findings of CSR is that it seems to provide a good framework for showing how reformed epistemology converges (remarkably or not) with the *attribution account* of CSR that we explore in the previous chapter. As mentioned there, several prominent cognitive scientists of religion attibute to our agency detection device a major role in the formation of religious beliefs. On the basis of experimental results (some of which we discusssed in the previous chapter), they claim that humans are "eager" – perhaps even "hyper"— detectors of agents. We are prone to identify objects and entities as agents on scant evidence. Even the slightlest evidence of the presence of agents is enough to trigger belief that there is or at least there might be another agent around. This propensity to detect agents even when their presence is ambiguous or potentially hidden has led cognitive scientists of religion to say that this is the same mechanism responsible for the human proclivity to believe in supernatural agents. And, according to M&D, this agency detection mechanism fits very well with their account of the *sensus divinitatis*. They claim that the *sensus divinitatis is*, as a matter of fact, "a sub-function of our agency detection device," "a prime example of our ability to grasp tacit support relations" (McAllister and Dougherty, 2018, 14), and that it serves to confirm their model.

If this is correct, we have in M&D's model an excellent example of how one of the three CSR accounts can be understood in a way that is consistent with evidentialism. It is time, then, to turn again to those three accounts and see whether what we have seen thus far provides us with what is necessary to assess the first claim, namely, that cognitive science (and CSR in particular) shows us that evidentialism is untanable. But before bringing this section to a close, it is important to take notice of the implications of what we have seen so far in this

chapter for evidentialism. All three models of the *sensus divinitatis* present themselves as consistent with a particular view of evidential support. Proper functionalists claim that the first model is an expression of their understanding of evidential support. Tucker's and D&M's models, on the other hand, are attempts to model the *sensus divinitatis* in ways that are consistent with the standard evidentialist approach to evidential support with its committment to *necessity*. Both Tucker's model, grounded in the phenomenal conservative understanding of evidential possession and evidential support, and D&M's model, with its focus on the mechanism of tacit perception of evidential relations, seem to make good progress towards showing how reformed epistemology and evidentialism can converge. Their models seem to secure an evidentialist understanding of the attribution account. Can they, or other varieties of evidentialism, secure the same result with respect to the dispositional and preparedness accounts? That's what we are going to explore in the next section.

To conclude our dicussion of reformed epistemology, there doesn't seem to be anything about it and about the *sensus divinitatis*, in particular, that precludes an internalist and evidentialist understanding of these theses. In fact, one would be hard-pressed to find a contemporary theist religious epistemologist (and even historical one, with the possible exception of Locke and the Lockeans) who would deny that one of the typical ways in which religious believers acquire their beliefs justifiedly is via some sort of non-inferential method. Besides Tucker and McAllister and Dougherty, other evidentialists who have written on religious epistemlogy, such as Paul Moser and Richard Swinburne, defend views that are clearly in accordance with the main tenets of reformed epistemology, as defined by Moon and Bergmann. Moser (2008, 2010), in fact a proponent of classical evidentialism, defends the view that certain religious beliefs can be justified on the basis of non-propositional evidence consisting of one's direct acquaintance with God himself, through the manifestation of his reality via our consciousness. And Richard Swinburne (2010, 2018), the paradigmatic natural theologian of our time, is a proponent of a variant of epistemological conservatism according to which religious beliefs can be justified on the basis of properly-basic beliefs and inclinations to believe that are themselves evidence. Evidentialism seems, therefore, to be at home with reformed epistemology. But can evidentialism be at home

with the findings of cognitive science? With the incredible contributions that cognitive science has been providing to our understanding of the human mind and, in particular, to our understanding of religious thought and behavior, the incompatibility of evidentialism with those findings would be a serious setback to that epistemological theory. Can religious evidentialim survive the rise of cognitive science of religion? That's what the remainder of this chapter will explore.

2. Evidentialism and the Three CSR Accounts

In the first chapter, we reviewed the current literature on the nature of evidence and saw that several of the most prominent proponents of evidentialism today favor a view about the nature of evidence according to which evidence consist of mental states. Among the most common ways of characterizing evidence in terms of mental states are: non-propositional or non-factive experiences, seemings, beliefs, inclinations to believe, dispositions, and intuitions. Let us take a closer look at each of these conceptions of evidence before we explore whether what the three CSR accounts have to say about the way religious beliefs are formed preclude the idea that subjects are responding in some way to evidence.

Several authors (Conne and Feldman, 2008, Moser, 1989, and McCain, 2014, to name a few) favor the view that evidence consists in non-propositional or non-factive experiences. Others, such as Dougherty (2011), favor the view that while experiences are the conveyers of evidence, the evidence itself is the propositional content of the experiences. The crucial role played by experiences in those accounts has led Dougherty and Rysiew (2013) to call them "experience first" views, in contrast to Williamson's (2000) knowledge-first epistemology, in that

> [experiences] are evidential regress stoppers. Any chain of cited evidence must end with the way the world appears to us to be. So on this view, experience is first in that it inhabits the ground floor of the intellectual edifice (Dougherty and Rysiew, 2013, 17).

Now, Moser and McCain are, as discussed in chapter two, explanationists (and Conee and Feldman seem strongly inclined to accept explanationism as well). They believe evidence

consists ultimately in non-factive mental states (and non-propositional in the case of Moser) and that the mental states that count as evidence possessed by the subject are occurrent mental states or (in the case of McCain's version of evidence possession) "mental states [one] is disposed to bring to mind when reflecting on the question of *p*'s truth" (2014b, 53). And those non-factive mental states become *justifying* evidence only when they figure in proper explanatory relations with propositions, namely, and roughly, only when the target proposition is part of the best explanation of those mental states.

One may think that this account places too heavy a burden on justification, restricting the glories that accompany the formation of justified beliefs only to the most reflective of us. But that doesn't seem to be the case. As mentioned in chapter two, on McCain's account, evidential support requires that evidence be available for *S* in *t* as part of the best explanation for why *S* has *e*. The criterion of availability is met when "at *t S* has the concepts required to understand *p* and *S* is disposed to have a seeming that *p* is part of the best answer to the question 'Why does *S* have *e*?' on the basis of reflection alone" (McCain, 2014b, 66). This theory doesn't require that the subject has a clear understanding of the concepts of "explanation" or "evidence." In fact, as a mentalist theory of justification, McCain's explanationaism doesn't include any awareness requirement on justification, but requires solely that the subject has a disposition, by virtue of her background evidence, to have a seeming that *p* is part of the best answer to the question above. What we have here is an approach to evidence possession and evidence support that relies on explanatory considerations but that could potentially be satisfied even by children and unreflective adults.[50]

Others (*e.g.*, Tucker, 2011) have favored a formulation of evidence possession in terms of seemings (see *PC* above) and that the degree of justification of a belief varies proportionally to the strength of those seemings, in the absence of defeaters. What is most characteristic of seemings is that they come with "forcefulness" (Huemer, 2001, 77), "assertiveness" (Tucker, 2010, 530), or "felt veridicality" (Tolhurst, 1998), and this "feel of a state whose

[50] McCain notes that children and unreflective adults may only have one explanation available to them and that they will have the disposition to have the seemng that the proposition in question is part of the best explanation because they will be disposed to have a seeming that the proposition is part of *the* explanation for the question "Why do I have *e*?" (2014b, 51%, endnote 34).

content reveals how things really are" (Tolhurst, 1998, 298-299) is plausibly what lead many to think that seemings confer immediate justification for their content (McAllister, 2017). Again, there is dispute over whether seemings are beliefs, inclinations to belief, or experiences, but it seems fair to say that the vast majority of defenders of the seemings view of evidence possession hold that seemings consist of experiences. According to this view, seemings have propositional content and are the sort of things that have a phenomenal character, though, apparently, not necessarily *qualia* (see the debate between Tooley, 2013, and Huemer, 2013b, on this point). Since phenomenal conservative is normally presented as an iternalist theory, its proponents are likely to regard seemings as necessarily occurrent (and that may spell trouble for phenomenal conservatives with respect to the justification of stored beliefs – see Tucker, 2013, 9), though there may be room for claiming that there could be non-occurrent seemings (see Bergmann, 2013, 160-161).

Others (*e.g.*, Lycan, 1998, and Swinburne, 2001) have defended the view that "the mere fact of believing a proposition confers some justification on that proposition" (Vahid, 2004, 102). This view is known as "epistemic conservatism" (though Huemer, 2013a, favors "doxastic conservatism," and Swinburne, 2001, prefers the expression "doxastic foundationalism"). Several motivations have been presented for this view: that the stability of one's belief system is a cognitive virtue; that it is unreasonable to change one's belief system without good reasons; that its rejection put us in the hand of the skeptic; that it is irrealistic to suppose that we could do otherwise, and so on. With respect to the last two considerations, Jonathan Kvanvig remarks that "any alternative to skepticism which enjoys the image of a piecemeal revision of our raft of beliefs will presuppose that there is some presumption in favor of the parts of the raft not under revision" (1989, 149).

Epistemic conservatism comes in many varieties (see Vahid, 2004, for a survey of these views). One way to express it is through the Principle of Credulity, whose defenders today include Lycan (1998) and Swinburne (2001) (but which can be traced back to at least to Thomas Reid, who called it the Principle of Testimony – see Swinburne, 2001, 141, footnote 14). According to Swinburne's formulation of this principle,

> This is the Principle of Credulity; the rational person is the credulous person; he is right to believe everything he believes as

> strongly as he believes it until it is rendered improbable by something else he believes. The Principle of Credulity says that the subject should start from where he or she is—doxastically (2011, 202).

According to Swinburne, our evidence consists of our basic beliefs. He regards the view that evidence consists of experiences – which is the most popular view among evidentialists today – as untenable, as other elements of our mental life ("purposes, desires, occurrent thoughts, and sensations" (2001, 203)) cannot be used "as evidence for anything else, except in virtue of what they believe (true or false) about them" (Ibid.). Thus, the internalist evidentialist should, according to Swinurne, "construe a subject's evidence as his basic propositions, that is mainly his basic-beliefs; and only his basic propositions" (Ibid.). But the internalist evidentialis should, on this picture, also construe available evidence as consisting of non-occurrent basic-beliefs that are readily available, beliefs that, according to Swinburne, "we are (subconsciously) 'aware of' them (they are not merely 'accessible') even when we are not currently thinking about them" (Ibid., 204). These are beliefs that, though we are not currently thinking about them, they are actively influencing the other beliefs we form and the actions we undertake. For instance, one can go to the refrigerator to grab some food while the thought that there was food in the refrigerator did not become part of one's conscious thinking. He seems to think that a mental state can only be properly called a belief one possess if it is involved (even if subconsciously) in how she forms other beliefs and acts. And these beliefs he regards as mental states we are "aware of," though not at a conscious level. These are "opertative beliefs," mental states that are readily brought to our attention if we are asked some question to which they are relevant. These beliefs, like those we are currently thinking, are available evidence, on this view.

The idea that believing *p* is evidence for *p* may sound counterintuitive, but many epistemologists (including those mentioned above) believe that there is no plausible alternative if we are to escape sckepticism. And at least one recent version of the principle, as advocated by McCain (2008), has been said to resolve a series of epistemological problems, such as the justification of memory beliefs, the problem of forgotten evidence, the problem of easy knowledge, and certain forms of skepticism. In any case, the crucial

point to draw from this discussion on epistemic conservatism is that thare are capable defenses of the idea that beliefs can be evidence. The mere fact that we hold them suffices, according to the proponents of this view, to provide those beliefs with some justification. And, according to this view, we should continue to hold those beliefs unless we acquire reasons to think that there are problems with them. Recall that this view that equates belief with justifying evidence is also regarded as a possible contender for the condition of right account of the nature of seemings. However, in light of Muller-Lyer cases and other illusions, this equation of the belief-as-justifying-evidence view with seemings seems to most participants in the debate over the viability of phenomenal conservative today as untenable (see, *e.g.*, Tucker, 2011, Huemer, 2013a and McAllister, 2017).

Others (*e.g.*, Sosa, 1998, and Swinburne, 2001, 2011) have suggested that inclinations to believe should be considered as evidence. Unlike the formulation of Lycan and others of the belief-as-evidence view, Swinburne's include not only beliefs, but also inclinations to believe as confering some justification for the proposition believed or to which one is inclined to believe. And that justification is proportional to the strength in which the belief is held or in which one is inclined to believe. As Swinburne put it, "[The] stronger the conviction that it is so forced, the stronger the reason for me to believe it to be true. If that were not so, I would never have good reason (in the sense of a reason of which I am aware) for believing anything" (2001, 140-141). He illustrates this latter claim with examples, one of them involving memory beliefs, which cannot be held on the basis of other beliefs. The phenomenology that accompanies or even characterizes our memory of past events will only constitute evidence about our past, Swinburne claims, if we form, or are inclined to form, beliefs about what happened in the past. In latter work, Swinburne seems to place an even stronger emphasis on the connection between inclinations and evidence (Swinburne, 2011). In fact, Conee and Feldman (2011, 294) claim that, in correspondence, Swinburne stated to them that "our evidence consists ultimately in [. . .] inclinations to believe, whether they are strong inclinations or weak ones" (Conee and Feldman's paraphrase). These inclinations are mental states that are "forced upon us by how things are in the world (whether or not we have a sensory awareness of what it is about the world which forces the belief upon us)" (2011, 201). They constitute our basic propositions, the "only way to get from our mental life

to the world" (Ibid). Among other considerations in favor of this view is that Inclinations to believe can contribute to render some beliefs probable. He provides the following illustration:

> If, as I watch the cars passing my window, I catch glimpses of several cars each of which it looks as if it might be red, although more probably it is black, and on the basis of each glimpse I ascribe to the proposition that the car in question was red the probability of ⅓, these inclinations together plausibly make it probable overall (more probable than not), in the absence of other evidence, that on at least one occasion a red car passed my window (Ibid).

Thus, according to Swinburne, inclinations seem to be the sort of thing that can provide evidence for propositions.[51] This view has been discussed in the literature on seemings as a potential candidate for the mental states that best characterize seemings.[52] Suppose, however, that seemings are really *sui generis* experiences with propositional content and characterized by the "feeling of truth" or of "forcefulness." Should we then say that the mental states refered to as "inclinations" that we have been exploring here do not exist? Or perhaps should we say that there are other mental states besides seemings with propositional content and that are characterized by that "feeling of truth" or of "forcefulness" that inclines us to believe their propositional contents?

It is important to note that some authors (*e.g.*, Huemer, 2013a, 2013b, Tucker, 2011, Werner, 2013) use the terms inclinations and dispositions interchangeably when speaking of the inclinations view of seemings. But while inclinations are occurrent mental states in that it is a mental state that is before one's mind, of which the subject is aware, dispositions are non-occurrent mental states, of which the subject is not aware at the moment, but that

[51] That is not to suggest that, on Swinburne's view, inclinations or beliefs provide, in and of themselves, justification good enough for justified belief. Swinburne defends the view that a belief is justified when it is made probable by the evidence and have adequate grounds, with probability understood here in terms of logical probability. And "the natural level to suggest for the adequacy of its grounds to make the belief justified is that the grounds give it probability of greater than ½" (2001, 56).

[52] Swinburne himself seems to equate inclinations with seemings (2011, 202).

could be brought to mind (perhaps easily) and that would pull the subject in the direction of believing *p* (see Swinburne, 2001, 140, note 13, for a similar distinction).[53]

So, there seems to be a distinct category of entities that can possibly count as evidence, entities that are distinct from inclinations. In fact, we have seen that McCain's explanationism, as described above, requires that the subject has a diposition to have a seeming that *p* is part of the best answer to the question "Why does *S* have *e*?" The role played by the disposition here is not, however, of ultimate evidence, but of "intermediate evidence that links *p* to *e*" (2014b, 81, endnote 27). The role of ultimate evidence is played, in his account of evidential fit, by "the background evidence that sustains this disposition to have the relevant seeming" (Ibid.)

There are good reasons, however, to think that dispositions themselves should be included in the category of ultimate evidence. The problem of stored beliefs for evidentialism that we discussed in chapter two clearly points in that direction. Recall from that discussion that most of our beliefs, even justified beliefs that seem to constitute knowledge, are non-occurrent. When we sleep and are not dreaming, all our beliefs are stored and we, supposedly, do not lose knowledge (for example, to repeat the example of chapter two, that the principle of noncontradiction is true) only because we are not awake (Moon, 2012b). However, it seems that the only source of evidence for these beliefs that could satisfy the evidentialist conditions on justification would be dispositions.

This has led evidentialists to embrace dispositionalism about evidence and justification, which can be expressed by the following schema:

> If *S* has disposition *X*, then *S* is prima facie justified in believing that *p* (Frise, 2017b).

[53] Werner (2013) prefers to make a distinction between two types of dispositions: those with phenomenal character and those without it. I think, however, that given the way "inclinations" and "dispositions" have been conflated in the literature, it is more helpful to distinguish the two. What Werner calls "dispositions" with phenomenal character is what is frequently refered to in the literature as the candidates for seemings known as "inclinations." So we should keep this terminology and refer to a distinct mental state that is non-occurent and, hence, without phenomenal character, as "dispositions." See Schwitzgebel (2002, 250) for statement that the typical understanding of dispositions in the literature is that they do not possess phenomenology.

But what exactly are dispositions? How do we acquire them? Are they always non-occurrent? Are they information that were at some point learned and then stored in memory? There are different answers to these questions in the literature. As mentioned above, the typical understanding of dispositions in the literature is that they are mental states that do not possess phenomenology (Schwitzgebel, 2002, 250). They can justify even though they are not manifest (Frise, 2017b). To say that S has an epistemic disposition is perhaps to say that *S* can recollect *p* or that *S* has the potential to bring *p* to mind. According to Conee and Feldman (2011, 304), dispositions justify the stored belief *p* when there is disposition to recollect *p* as known.[54] If one has a disposition to recollect *p* as known, one has evidence that justifies the stored belief that *p*. And in order for *S* to have a disposition to recollect *p*, *S* must have learned and not forgotten *p*. However, Conee and Feldman seem to understand "learned" in a broad sense of the word, for dispositions to recollect can originate, according to them, from "brain malfunctioning or tempering" and can *prima facie* justify "whatever its historical origins" (2011, 305).

It is important to distinguish, as does Audi (1994), dispositions to believe from dispositional beliefs. The former are dispositions, whereas the latter are (stored) beliefs. According to Huemer (1999, 356, footnote 15), stored beliefs must have been at some point occurrent. Frise, however, rejects this view:

> there is reason to deny that a stored belief must at some time have been occurrent. Information that enters memory is normally altered in at least three stages of memory processing. If memory stores nearly as many beliefs as we think it does, we should allow that we believe the content resulting from this processing. But much of this content has never been occurrently endorsed. So a stored belief need not have been occurrent at any time (Frise, 2017a, 491).

[54] Frise suggests that it is better to understand the justification of recollective dispositions more broadly, including not only dispositions to recall *p* as known, but also as true, and the experienced events and set of propositions as jointly and clearly indicating that *p* (Frise, 2017b). In the end, however, he finds that even this modified version of Conee and Feldman's dispositionalism is unsatisfactory.

The three stages of alteration of memory mentioned by Frise are: encoding, consolidation, and retrieval. In the first stage, the information that enters into memory is initially stored. But this initial process of storage does not occur passively, however. Rather, a process of content alteration of the information begins, with the "gist" of the information being extracted and stored. The second state continues with content alteration, but now new content (about general features of the information) is also generated. In the third and final stage, the retrival of the information is constrained by the context, so that what is recalled only partially resembles what was initially stored. This context involves the cues that prompt recollection, the related information that was recently retrived, and the constitutive elements of the retrived content. The first contextual factor refers to the fact that the quantity and quality of the retrived information varies according to the manner in which the information is cued. The right incentives can bring out greater quantity and greater quality of information. The second contextual factor has to do with how information previously retrived can affect the quantity and quality of the information extracted in the next round of retrival. And, finally, the third factor involves the process known as *source monitoring*, which help us discern whether the retrived information originated from what we experienced or from what was merely imagined.

In sum, on this standard picture of the functioning of memory that has emerged after a few decades of experimental research, information is altered at all three stages, with the final result, at retrival, looking a lot different from what entered one's memory at the beginning of the first stage. This process is constructive from the beginning to the end. Yet, "interestingly, none of this discredits memory" (Frise, 2018, 67). While differing in important ways from the original information, the belief formed at retrival tends to be true. Source monitoring, which indicates to us whether the origins of the retrived information are to be found in our experience or our imagination, plays a major role in securing the reliability of memory. Frise summarizes this process as follows:

> memory dismantles and highly selectively stores bits of inputs, and then adds to these inputs, and then adjusts them in light of features of the retrieval context, and only then does a recollection get its exact content, and nothing about memory by itself

> guarantees that this content will even be endorsed; whether we endorse what we retrieve is determined by our confidence, at the time, in what is retrieved. And a metamemorial process uses several factors from the retrieval context to produce this confidence. Recollected content is best explained as generated (Ibid.).

This picture of memory (again, the standard picture that has emerged from research in cognitive psychology) spells trouble for representationism about belief. Representationalism is at home in a picture of memory that resembles a storehouse, where beliefs – *i.e.*, mental representations – are stored until they are retrived. On representationalism, beliefs are mental representations. The problem that cognitive psychology poses for representationalism is that, as we have just seen, beliefs are not simply stored in memory – they are "dismantled," reconfigured, and reassembled, emerging from recollection bearing little resemblance with the original belief. And yet, the retrived belief tends to be true and the process reliable. The core information is stored, but not the mental representations. If representationalism is true, then there are no stored beliefs. Our memory generates beliefs; it doesn't preserve them.

On the other hand, if we take beliefs to be sets of dispositions, or, more precisely, "a suitable combination of dispositions pertaining to *p*" (Ibid., 69), then we can say that there are in fact stored beliefs. For, it doesn't matter, on this picture, whether mental representations are filtered and dismantled. What matters is whether we will have the suitable dispositions pertaining to *p*. What we have here is a potential solution for the problem of stored beliefs for evidentialism.[55] Recall that the problem is, roughly, one of identification of the source of evidence for these beliefs that could satisfy the evidentialist conditions for justification. On Frise's suggested dispositionalist condition on justification, a stored belief that *p* is justified when "S has a suitable combination of dispositions pertaining to *p*" (70). To use one of Frise' examples: when your cellphone rings and your automatic response is reach to your pocket

[55] In fact, Frise argues that on both accounts of the nature of belief the problem of stored beliefs is solved: if representationalism is true, there are no stored beliefs, and, therefore, no problem of stored belief; if dispositionalism is true, we have a plausible account of how stored beliefs can be justified – if S has a suitable combination of dispositions pertaining to *p*, S's stored belief that p can be justified.

to pick it up, you weren't thinking about the location of your cellphone. You non-occurrently believed your cellphone was in your pocket because you had the suitable combination of dispositions. And that belief was justified. Even before your cellphone rang, you were disposed to have a recollecting experience of the location of your cellphone and feel confident about it. Hence, on this picture, even when one is not thinking about *p*, one may be justified in her belief that p, provided she has the right set of dispositions pertaining to *p*.

Thus, to sum up, "evidentialists need not appeal only to conscious occurrences as justification" (Conee and Feldman, 2011, 304). They need not appeal only to dispositions that constituted at some point information learned through a conscious process, *i.e.*, that entered the mental life of the subject reflectively before they were stored. And they need not appeal only to information that was stored exactly as they entered the subject's mental life and that didn't go through any process of filtering and modification. Evidentialists have, therefore, a considerable range of possibilities in how to understand evidence and the role evidence plays in justification.

Ryan Byerly (2014) has developed a general account of epistemic support in terms of dispositions similar to Frise's proposal for memorial justification (2018, 70, 71). On Byerly's theory, all that is needed for justification (memorial or otherwise) is that the subject be disposed to take a doxastic attitude toward *p* in light of her total evidence:

> *S* is justified in believing *p* when *S* is disposed, in light of her total evidence, to believe *p*;
>
> *S* is justified in disbelieving *p* when *S* is disposed, in light of her total evidence, to believe non-*p*;
>
> *S* is justified in suspending judgement about *p* when *S* is disposed, in light of her total evidence, to suspend judgement about *p* (2014, 405).

The main motivation for this account is that, while it is an internalist theory of epistemic justification (a disposition is, after all, the sort of thing that is in one's mind)[56], it avoids what

[56] And supposedly, make a contribution to one's epistemic life. And, thus, subjects who seem to be mentally alike, except that one of them lacks the relevant dispositions, are not mentally alike after all. See the debate between Moon (2012a, 2015) and McCain (2014a, 2014b) on this point.

Byerly perceives as the overintellectualization of epistemic support by the alternative evidentialist views, which "imply either that nobody's evidence ever supports a proposition or they threaten to imply that only the more sophisticated among us are such as to have evidence that supports a proposition" (2014, 410). On Byerly's theory, no predisposition to have a seeming about the best explaation of one's experiences is required. Nor is direct acquaintance with evidential relations, or the satisfaction of the axioms of the probability calculus, and so on. Believing (or suspending judgement) according to what one is disposed to believe (or suspend judgment) in light one's total evidence, suffices for justification. As discussed in chapter two, Byerly (2013) argued that explanationist evidentialism cannot account for the justification of certain beliefs about the future. In one of his examples, *S* gives the last shot in a last round of a golf game in which he has performed well. The hole is very close and the ball is going in the right direction and speed. *S* forms the belief that the ball will fall into the hole. Byerly claimed that McCain's explanationism could not deliver the result that this belief is justified. On the other hand, his dispositionalist evidentialism has no problem with such a result, for, as long as *S* is disposed, in light of her total evidence, to believe that the ball will fall into the hole, that belief is justified (2014, 407). Another alledged probem with explanationism not shared by Byerly's dispositionalism is that it is enigmatic why, when the realization conditions of the disposition are satisfied, S does not have an occurrent seeming. Suppose that *S* is disposed to form the belief that her social security number is 343-43-0974. On explanationism, *S* may be justified if she is disposed to have a seeming that that's the best explanation for her memorial experience. However, why would *S* only be disposed to have a seeming, instead of having an occurrent seeming, if that's what she is disposed to in light of her total evidence? (2014, 416).

And, finally, intuition is a concept prolifically employed by cognitive scientists. The dispositional CSR account that we will revisit below is, to a large extent, grounded on the idea that a major factor in making religion so widespread is that some religious concepts are minimally counter*intuitive*, and such concepts are more attention demanding and, hence, more easily remembered and tramsmissible. These cognitive scientists seem to think of intuitions as "expectations" and "assumptions" (Barrett, 2008) that emerge from "implicit proceses in the mental basement" (Boyer, 2001, 305) that are distinct from beliefs but that

nonetheless affect belief formation (though, in some contexts, they seem to use “intuition” as a kind of belief – Audi, 2014, 18). Intuitions form the basis of our folk physics, folk biology, and folk psychology, on this view. Counterintuitive concepts or ideas are those that challenge our natural way to think about the physical, biological, and psychological aspects of reality. Thus, minimally counterintuitive concepts counter our intuitions, our basic ontological categories, our natural way to see the world and its constitutive entities, just enough to become attention demanding and easily retained and transmissible (see, *e.g.*, Barrett, 2004, chapter 2).

Like in cognitive science, philosophical talk of intuitions is far from monolithic. Nonetheless, overall, intuitions are regarded by philosophers as a specific sort of seemings, distinct from memorial, perceptual, and inferential seemings. As Michael Huemer put it,

> We form justified beliefs about the physical world on the basis of sensory appearances (also called 'perceptual experiences'). Similarly, we form justified beliefs about certain abstract truths, including [. . .] self-evident principles [. . .], on the basis of intellectual appearances ('intuitions'). Intuitions are mental states in which something appears to be the case upon intellectual reflection (as opposed to perception, memory, or introspection), prior to argument (2005, 232).

Consequently, on this, predominant view of intuitions, they convey truths directly, not as the result of inferences, not by the five senses, and are mental states that normally arise prior to belief (see Huemer, 2005, 6, 10). This seems consistent with the way cognitive scientists, such as those cited above, see intuitions. One possible point of disagreement among most philosophers writing on intuitions and cognitive scientists, however, is that, unlike the latter group, the former tend to reject the idea that intuitions can in some contexts be some sort of unconscious belief (see Huemer, 2005, 103; Audi, 2014). What is important for our purposes here, however, is whether intuitions can be evidence. The philosophical debate on whether intuitions are evidence and can justify beliefs is a burgeoning field of study today, attracting the interdisciplinary contribution of cognitive scientists and experimental psychologists, which has led to a plethora of experimental challenges to the reliability of our

intuitions (see Pust, 2017, for a survey of the literature). One thing that seems clear, though, is that philosophers do use intuitions as evidence in their philosophical work. Climenhaga (2017) has argued that philosophers do take intuitions to be evidence on the basis of three considerations: philosophers tacitly rely on intuitions as evidence, they offer error theories for intuitions that they reject, and they are more confident in accepting theories when they have intuitions about diverse cases supporting them (and conversely). That philosophers use intuitions as evidence doesn't mean that they are evidence, of course. Philosophers who are skeptical of the idea that intuitions provide evidence are usually those who reject the view about the nature of evidence according to which experiences constitute evidence (Williamson, 2000, 2007). On the other hand, philosophers who accept that seemings provide evidence – which seems to constitute the vast majority of the participants in this debate – have, in principle, no problem accepting the view that intuitions are evidence.

Each of these views that we have examined in this section faces objections, to be sure. But since our goal here is not to defend the truth of any of these particular views, but simply to present some varieties of plausible accounts of the nature of evidence, evidence possession, and evidentialist epistemic justification that will allow us to evaluate *claim 1*, we need not defend those views from objections or attempt to show that the objections succeed. It's time now to explore whether religious beliefs, as formed according to the mechanisms of the three CSR accounts discussed in the previous chapter, can plausibly be understood as involving evidence as conceived by those accounts, and perhaps as being justified in terms of those evidentialist accounts of justification.

2.1. Attribution Account

We have now what we need for our exploration of whether evidentialism is compatible with the findings of the three CSR accounts. With respect to the first account, we have seen that Tucker and McAllister and Dougherty have developed evidentialist models of the *sensus divinitatis*.[57] On the first of these models, the *sensus divinitatis* is seen as turning experiences

[57] With respect to the compatibility of Calvin's and Plantinga's model with this account, Clark and Barrett say that: "in terms of input conditions, [the attribution account] looks more like Plantinga's experientially-triggered faculty. Especially according to Bering's version that allows for a broader set of triggering experiences, this god-faculty generates empirically grounded beliefs in gods through means similar to Plantinga's characterization.

such as visual sensations into seemings, which constitute evidence for religious beliefs. On the second model, the *sensus divinitatis* is seen as consisting of our standard ability to tacitly perceive evidential relations.[58] Recall that McAllister and Dougherty claim that their model is consistent with what CSR has been telling us about supernatural agency detection.

According to the attribution account, humans are "eager detectors of agents." A series of experiments (some of which were discussed in the previous chapter) have shown that newborns are able to discriminate between the biological and the non-biological patterns of motion and attend preferably to biological motion and that that seems to indicate that human beings are already predisposed to detect agency since their first two days of life. Infants also attribute agency to ambiguous stimuli. Three-month-old infants attribute agency to a self-propelled inanimate object when cues are given that the object is an intentional agent. And when they see an inanimate object moving without the obvious presence of an agent, they expect the presence of a hidden agent.

In fact, we not uncommongly find ourselves easily identifying objects and entities as agents on scant evidence. A noise or shadow in the middle of night, or a particular formation of a cloud, can be enough to trigger belief that there is a burglar trying to break into the house, or that there is a face in the cloud. For this reason, this mechanism has been termed "hypersensitive." We are hypersensitive agency detectors. Even the slightest evidence of the presence of agents are enough to trigger belief that there is or at least there might be another agent around. A natural explanation for the origins of such cognitive mechanism, now called by many "hypersensitive agency detection device" (HADD), is that it provided an evolutionary advantage to our ancestors who possessed them. For, as the saying goes, "better to err on the side of caution."Hypersensitive detectors of agencies had the evolutionary advantage of being less likely to be caught by surprise by predators or enemies. This evolutionary advantage conferred by the capacity to detect agents on scant evidence has been considered by many cognitive scientists of religion as being one of the main

Considering this god-faculty's outputs, however, it lacks the specificity of Plantinga's and looks more like Calvin's general, hazy sense of superhuman agency of some sort" (2010, 179).

[58] On this model, the *sensus divinitatis* is not a single God-faculty, but "a tendency to draw connections between the content of our experiences and propositions implying the existence of God" (McAllister and Dougherty, 2018, 12).

mechanisms involves in the formation of religious belief, and, thus, largely responsible for the ubiquity of religion.

Of all the three main CSR accounts on the primary mechanisms that produce religious belief, the attribution approach is the one most clearly ameanable to an evidentialist interpretation. This is due to the fact that this approach is the only one that takes perceptual experience to be the trigger of religious belief. Naturally, the two evidentialist accounts of the *sensus divinitatis* discussed in the first section can easily accommodate the attribution account. Tucker's model will say that the perceptual experiences involved in agency detection will trigger the *sensus divinitatis*, which will, in turn, produce a seeming that will constitute the evidential basis of the religious belief. McAllister and Dougherty's model, on the other hand, say that the perceptual experiences involved in agency detection will produce a seeming, which, in conjunction with the subject's background information, will trigger the *sensus divinitatis* (2018, 12-13). So, the agency detection mechanism fits very well with the M&D model of the *sensus divinitatis*. They claim that the *sensus divinitatis is*, as a matter of fact, "a sub-function of our agency detection device," "a prime example of our ability to grasp tacit support relations" (14), and that it serves to confirm their model.[59]

Can the CSR findings on supernatural agency detection be accommodated into the other evidentialist frameworks that we have explored in the previous sub-section? The short answer is that the attribution account seems, in fact, to be consistent with all the evidentialist theories discussed above. Let's say something about a few of them:

[59] It is important to note that M&D only address the question of the compatibility between their model and CSR with respect to the attribution approach, and they boldly claim that the hypersensitive agency detection device (HADD) "is arguably the most prominent theory in the cognitive science of religion"(2018, 14). But while HADD may perhaps be the most widely *discussed* CSR theory in the literature today, it is still a somewhat contentious theory. Pascal Boyer (2001), for instance, is skeptical that HADD plays a major role in religious belief. He believes that HADD mistakes are easily overridden and that even when beliefs are formed with the participation of this mechanism, they are normally unstable. And many who accept the theory, such as Barrett and Lenman (2008), do so with reservations: "*We are not arguing that HADD experiences are directly responsible for belief in supernatural agents.* We are arguing that HADD experiences, belief in MCI [minimally counterintuitive] agents and discourse about such agents are *mutually reinforcing*. *HADD experiences can help encourage, reinforce, and spread belief in MCI agents*. For example, having a HADD experience with no obvious natural explanation in a location that one has just been told is the site of frequent divine appearances will make belief in those appearances more plausible. Similarly, exposure to discourse about MCI agents or having a reflective belief in them can increase HADD experiences, as hearing about such ghosts or gods increases the HADD's vigilance" (2008, 116, my italics).

- Epistemic or doxastic conservatism says, roughly, that "the mere fact of believing a proposition confers some justification on that proposition" (Vahid, 2004, 102). So, the conservative has a straightforward answer to the compatibility question: beliefs formed via the mechanisms of the attribution approach may have some justification;

- Proponents of the inclinations to belief view of evidence will say that inclinations to belief produced by the mechanisms of the attribution approach are evidence for the propositional content of those inclinations. And proponents of the inclinations view of seemings will say that if those mechanisms produce inclinations that are seemings, then the subject might have some justification for her religious beliefs formed on the basis of those seemings;

- Explanationism also seems capable of delivering the correct results with respect to the attribution account, as the mechanisms described by that account may trigger dispositions in us to have seemings that the supernatural agents that we believe we are detecting are the best explanation for our experiences;

- Dispositionalist, such as Byerly (and perhaps Frise), could perhaps say that God designed us in such a way that, as we develop cognitively, we would become disposed to believe certain religious propositions in light of our total evidence. In terms of the attribution approach, our perception of supernatural agents could make us disposed to believe that there is such-and-such supernatural agent in light of our total evidence and such a belief could be justified on this account.

In sum, the mechanisms described by the attribution approach seem compatible with a wide variety of evidentialist accounts of evidence, support, and justification.

2.2. Dispositional Account

Why are some concepts more commonly observed across cultures and throughout human history than others? Proponents of the CSR disposition approach answer this question by saying that some concepts are more successful in activating certain mental systems, which makes them more easily remembered and transmissible. Those conditions are, essentially, that the concept violates certain expectations about ontological categories (*i.e.*, they are

counterintuitive) while preserving other categories, and this makes them particularly attention-demanding and transmissible. For instance (to borrow one of Barrett's examples), "what captures your attention more, a potato that is brown, a potato that weighs two pounds, or an invisible potato?" (2008, 152).

In addition, such concepts must have rich inferential potential, that is to say, they must preserve default inferences that put the mind in the position to provide information that completes the fragmentary elements initially given, thus making the idea more memorable. For instance, a ghost is a person with counterintuitive physical properties. And if you believe you have seen a ghost "your mind creates a whole lot of assumptions of which you are not necessarily conscious. [. . .] You assume that the ghost saw you [. . .] knows [. . .] remember[s] . . ." (Boyer, 2001, 72).

If a concept is to become highly memorable and transmissible it is not enough that it is counterintuitive and with inferential potential. It must have cunterintuitiveness at the exact right level – it must be minimally counterintuitive (MCI) in order to become culturally and historically widespread. It must violate the expectations about ontological categories at just the right level to become attention demanding. Potatoes, for instance, meet our expectations about ontological categories. And once we add the idea that a potato weighs two pounds, some violation occurs, and the concept becomes to some extent counterintuitive, but with little inferential potential. When we add that the potato is invisible, that adds even more counterintuitiveness and makes for greater inferential potential – "where is the potatoe at this exact moment?" But if we say that there is a potatoe that ceases to exist after five seconds of existence, while it was a counterintuitive concept, it lacked any relevant inferential potential that would make it memorable and transmissible.

Some religious beliefs are just the sort of concepts that meet the requirements for minimal counterintuitiveness – the sort that has endured the test of history and become widespread across cultures. And because they are the sort of concepts that are attention-demanding and that produce substantial inferential potential, proponents of this CSR approach to how we form religious beliefs argue that we are relatively disposed to come to think about – and perhaps believe – these religious concepts. Thus, on the dispositional account, certain

religious ideas, in particular ideas about supernatural agents such as God or gods, meet certain conditions that make those ideas relatively easy to memorize and be transmitted. This is, according to proponents of the dispositional account, the primary reason why certain religious ideas are so prevalent and pervasive. God concepts are easily and readily represented by the human cognitive equipment, are attention-demanding, have rich inferential potential, and are important enough to motivate action and reinforce belief. Unlike what the attribution account tells us, here experiences are not needed to trigger belief. This model is epidemiological: the human mind is particularly susceptible to being infected by certain religious ideas. Once we understand those ideas, we are inclined to "embrace them, act upon them, and pass them along" (Clark and Barrett, 2010, 182). The inputs here are, therefore, not experiential, according to Clark and Barrett. The outputs of these dispositions, when it comes to god beliefs, on the other hand, are vague: "the god will not be identical with a human (or it would not be attention-demanding and have richer inferential potential), but it could have any number of properties such as being invisible, superknowing, superperceiving, superpowerful" (Ibid.).

Presumably, when one comes to believe certain religious propositions on the basis of this account alone, without the influence of other mechanisms or sources of belief-formation,[60] one is forming those beliefs on the basis of seemings, or inclinations to believe, or because one is disposed to believe those propositions. If any of these conditions is present when belief is formed on the basis of the mechanism described by the CSR dispositional approach, then evidence can be said to be present in the belief-formation process and one can plausibly claim that evidentialist conditions on justification are being satisfied. If such propositions seem true to us or we are inclined to believe them, or we are disposed to believe them, or (on doxastic conservatism) the mere fact that we hold beliefs about them, then a number of evidentialist conditions of justification (conservative, dogmatist, credulist, dispositionalist, and perhaps explanationist) may be satisfied when belief-formation involves

[60] This is a highly idealized picture, to be sure, for it is unlikely that many persons – if any – come to hold religious beliefs – in particular theistic beliefs – on the basis of one of these accounts in and of itself (and in particular of the CSR dispositional account), without the influence of other considerations and sources of belief-formation. But for the purposes of investigating whether each of these mechanisms could potentially pose a problem for religious evidentialism we are omitting this more realistic picture of belief-formation.

the mechanism described by this CSR account. But, again, as noted in the previous footnote, this is a highly idealized picture of formation of religious belief. It is unlikely that many individuals form beliefs exclusively on the basis of such an account, without the contribution of other sources of belief-formation.

2.3. Preparadness Account

As mentioned previously, several experiments have provided robust evidence that young children have a promiscuous tendency to view both non-living natural phenomena and living beings as existing for functions or purposes. One procedure that has been adopted to test children's teleological thinking is to invite them to play a game with the experimenter and two fictitious characters in photographs. In one of the experiments, the charecters ("Ben" and "Jane") were introduced as persons who "love to talk about different things but never ever agree with each other". The children were told to listen to what Ben and Jane had to say about a certain subject and then point to the character that they thought was right. One of the items presented to them was the following:

> See this. This is a tiger.
>
> Ben says a tiger is made for something. It could be that it's made for eating and walking and being seen at the zoo or it could be that it's made for other things. But Ben is sure that a tiger is made for something and that's why it's here. Jane says that this is silly. A tiger isn't made for anything. Even though it can eat and walk and be seen at the zoo, that's not what it's made for, they're just things it can do or people can do with it. Jane is sure that a tiger can do many things but they aren't what it's made for and they aren't why it's here.
>
> Point to who you think is right. Ben who thinks a tiger is made for something or Jane who thinks that's silly because a tiger isn't made for anything (Kelemen, 1999b, 256).

The children were as likely to say that the teleological explanation was the correct one as to favor the non-teleogical answer. This experiment setup was apparently designed to draw

from children their views on what was the best explanation for specific phenomena (the existence of tigers, birds, trees, rocks, and so on), and they consistently selected the teleological hypothesis as being the best. When asked why non-living natural phenomena such as pointy rocks exist, children often gave teleological answers, such as “so that animals will not sit on them” or “so that they can scratch their backs” (Kelemen, 1999a). And when asked why living beings such as birds exist, they often said things like: so that “they can fly” or so that “they can eat and grow” (1999b). From these findings, Kelemen (2004) suggested that children seem to be “intuitive theists,” in that they seem to be “predisposed to construe natural objects as though they are nonhuman artifacts, the products of nonhuman design” (295).

These results seem consistent with the accounts of justification advanced by both the dispositional and the explanationist varieties of evidentialism. With respect to the latter, the experimenter is in effect asking the children which of the two explanations offered by the characters for the content of the photopgraph is the best one. Experimenters have found tha children consistently choose the teleological answer. Recall that on McCain’s explanationism, one is propositionaly justified when the evidence is available for *S* at *t* as part of the best explanation for why *S* has *e*. And the criterion of availability is met when “at *t S* has the concepts required to understand *p* and *S* is disposed to have a seeming that *p* is part of the best answer to the question ‘Why does *S* have *e*?’ on the basis of reflection alone” (McCain, 2014b, 66). Thus, one of the implications of this account is that even children and unreflective adults can be justified in some of their religious (and perhaps theistic) beliefs. This may be the case even if they have only one explanation available to them. They can be justified in their religious belief as long as they have the disposition to have the seeming that the proposition in question is part of the best explanation for the question “Why do I have *e*?” simply by being disposed to have a seeming that the proposition is part of *the* explanation for the question (2014b, 51%, endnote 34).

Studies have shown that children not only can produce explanations, but sometimes even sophisticated ones, and that they, in fact, employ the same essential explanatory structure used by adults and even scientists. Keil and Wilson (2000) write that “even prelinguistic children” seem to possess "at least a rudimentary form of explanatory understanding” (4).

And, in their survey of the literature on children's explanatory reasoning, Brewer, Chinn and Samarapungavan (2000) report, among other things, that:

- there is some evidence that children prefer simpler theories;
- there is strong evidence that children favor theories that are consistent with their background beliefs;
- explanations children offer for physical phenomena display the same essential structure as those used by scientists, except that children do not offer explanations under the assumption that explanations should be testable;
- "children use most of the common forms of explanatory frameworks used by scientists, except for formal or mathematical accounts, and may use some additional forms that scientists do not use" (294);
- children behave as "little scientists," learning about the world and developing their own theories about the natural world, in a way similar to scientists;

They conclude from their review of the literature that

> qualitatively, children show competence with most aspects of everyday explanations at an early age. However, those aspects of explanation that derive from the social institution of science [. . .] are much later to develop and in some cases may never develop without the explicit training involved in becoming a member of the scientific community (296).

A less cautious conclusion is drawn by Alison Gopnik (2000). He claims that there is an almost complete overlap in the mechanisms children and scientists employ to acquire knowledge of the world. Nonetheless, "it is not that children are little scientists but that scientists are big children" (301). It's not just that children and scientists share similar explanatory mechanisms and strategies: science is built on these basic abilities, but organizing them in a larger and specialized social context, and applying them to new problems and domains. In addition, scientists are big children in the sense that, as McCauley (2000, 66-67) notes when commenting on Gopnik's views on this subject, actual scientists

exhibit cognitive limitations and biases also manifested by other humans, including by the "sophisticated little scientists."

In sum, these studies tell us that children display a far more sophisticated capacity to reason explanatorily than it is commonly acknowledged. And, as experiments such as those conducted by Kelemen and mentioned at the beginning of this subsection seem to demonstrate, even young children manifest religious beliefs about the creation of living and non-living things in circumstances involving inference to the best explanation. They do so because they seem to be prepared, *i.e.*, disposed to have seemings to the effect that those entities have an origin and have been created for a purpose. As discussed in the previous chapter, children do not anthropomorphise the creator, as Piaget thought. Rather, they are fully capable to reason abstractly about the creator. All this suggests that children can be justified in some of their religious (and perhaps theistic) beliefs on the explanationist account of justification. And if they can be justified on an explanationist account such as the one developed by McCain, they can clearly be justified in the less stringent dispositional and conservative (both phenomenal and doxastic) accounts. So its seems that evidentialism, in the varieties we have examined here, is well suited to account for the justification of particular religious beliefs as described by the mechanisms of the preparedness account.

Although only the previous CSR approach is label by Clark and Barrett (2010) "dispositional," the preparedness approach is likewise predicated on the idea that we have certain dispositions that will develop and mature as we develop cognitively. Both approaches, unlike the first one, are not perceptual. Recall that the reductive model of the *sensus divinitatis* posited the existence of a system of tacit perception of evidential relations as the underlying mechanism that produces theistic seemings and then beliefs. This model fits very well with the findings of the first CSR approach, which involves the detection of agency, but not so well with the second and third approaches. Both the dispositional and the preparadness approaches don't seem to be most obviously described as involving the sort of perceptual mechanism described by the reductive model. Rather, they involve dispositions that are triggered in the right circumstances or induced by the right cues and suggestions.

Thus, the reductive model with its focus on tacit perception of evidential relations may not be the most adequate to account for the way the dispositional and preparedness

approaches to CSR findings describe the development of religious thinking. If we want to model the *sensus divinitatis* in a way that is consistent with what cognitive scientists are saying about the cognitive origins of religious belief, we would do well to look for an evidentialist model that considers dispositions to be evidence and justifiers. That's what we are going to do in the fourth section. But before we do that, we need to examine the more general objections to evidentialism from cognitive science and show why they fail – if they fail.

3. Three Objections from Cognitive Science

From what we discussed in the previous section, and on the basis of the experiments and theories explored in the previous chapter, it sems clear that there is nothing in the three CSR approaches that threaten religious evidentialism. Each of those accounts can be interpreted in terms of religious evidentialism. It is too soon for the evidentialist to declare victory, however. For, while specific CSR accounts about the origins of religious belief may not pose any serious problem for evidentialism about the justification of religious beliefs, there may be objections that come from other considerations related to cognitive science that would call evidentialism into question. And, in fact, there are at least three objections in the literature that purport to show the inconsistency of the findings of cognitive science and evidentialism about epistemic justification. In this section we will present those three objections, reconstruct them in the form of valid arguments, and explore the strength of their premises. Perhaps once the premises are identified and they are examined in light of what we have seen thus about the nature of evidence and about the varieties of contemporary evidentialism, those arguments will be shown to rest on faulty foundations. As we will see, that's precisely the case.

4.2. Cognitive filtering and processing of information

Justin Barrett has defended the view that the picture of our cognitive architecture and processing of information advanced by cognitive scientists poses a serious problem for the evidentialist picture of justified belief formation. Barrett (2004 and 2009) compares the human mind to a workshop, where specialized tools contribute for the processing of specific classes of information. Those mental tools construct most of our beliefs. But they don't do

so by merely absorbing what is experienced, processing it, and storing it without biases and distortions. Rather, they have built-in biases that select, process, and store information in a manner that doesn't correspond to the picture of mind as a passive absorver and storer of information. As the information is processed, it is filtered according to its built-in biases and systematically distorted.

All of this occurs at an unconscious level. "The majority of these beliefs," notes Barrett, "are never consciously evaluated or verified" (2009, 79). These are what Barrett calls non-reflective beliefs. There are, nonetheless, beliefs that are reflectively formed. Even these beliefs, however, are subject to the selective activity of our cognitive system. In fact, reflective beliefs are constructed from non-refletive beliefs. As Barrett put it: "Specific mental tools generate non-reflective beliefs relevant to a given domain but then more general mental processes draw upon available non-reflective beliefs to form reflective beliefs" (80). Here is na example he uses to illustrate how these mechanisms and types of belief Interact:

> Some reflective beliefs arise from the converging outputs of several mental tools or multiple outputs of a single mental tool. For instance, if I observe a man with a bulge under his shirt trot out of a store, not pausing for a store worker standing in his path but running right over her, what am I to believe? My well-rehearsed human form detector tells me the bulge is not part of human anatomy (I don't even consider the man has a tumor). My Theory of Mind tells me automatically that a bulge under a shirt is not visibly accessible to others and because intentional agents behave purposefully, the man desired to obscure the object. As my Agency Detection Device knows that intentional agents need not continue on inertial paths, I non-reflectively assume the man did not intend to stop for the store worker. Theory of Mind tells me the man must have desired to run over the store worker and sends me searching for an explanation of this behavior (since, unlike inanimate objects, we don't typically attempt to move

> others through contact). The explicit suggestion that the man was stealing from the store makes sense intuitively given the outputs of all these mental tools working in concert. Hence, I am likely firmly to adopt this reflective belief even though I have no conclusive evidence. It just seems to make sense. In general, the more different non-reflective beliefs that converge on a particular reflective candidate belief the more likely the reflective belief becomes held (80-1).

In the end, then, even if we respond to the evidence available reflectively, we do so in a way that is never dissociated from the operations of the the mental tools and non-reflective beliefs. The environmental inputs that may be characterized as evidence are not preserved by our cognitive system: "The evidence (if available)," Barrett claims, "is always filtered and distorted by the operation of mental tools. We never have direct access to evidence but only processed evidence – memories" (81). Barrett concludes from this that there is no such a thing as "pure experience" (81), something he – and other critics of evidentialism from cognitive science, such as Greco, take to be part of the picture of belief formation pressuposed by evidentialism. No pure experience, no pure evidence. Evidentialism is in real trouble.

The environmental inputs that may be characterized as evidence are thus subjected to a process of filtering too unlike what evidentialism supposedly requires. Ultimately, reflective beliefs are not formed in response to "pure experience" and they are never spared from the influence of mental tools and their built-in biases. The evidence does not enter in our noetic system and is evaluated reflectively by our mind as it is out there in the world, objectively and in an unbiased way.

4.3. "Non-conscious" perception and sub-personal inputs

John Greco presents a similar objection in *Achieving Knowledge* (61-2). One of the problems he sees with evidentialism is that it "is undermined by contemporary cognitive science." He grounds this claim in the idea that "recent empirical studies make it doubtful that paradigm cases of knowledge, such as perceptual knowledge, memory knowledge, and inductive

knowledge, can be understood entirely in terms of person-level representational states, as evidence is understood to be." He cites blindsight and other kinds of "non-conscious" perception as examples of cases of knowledge that depend on sub-personal inputs that cannot be "appropriately understood as evidence." For Greco, blindsight and other kinds of "non-conscious" perception involve reliable faculties, but "their reliability is not entirely a function of person-level representations, such as perceptual appearances or 'seemings.'" Rather, their reliability depend on sub-personal inputs that do not sufficiently indicate the truth of *p* and, therefore, should not be considered evidence.

In addition to that, his objection to evidentialism from cognitive science is grounded on the Idea that

the reliability of our cognition seems to depend on "processing that operates on a sub-personal or even non-representational level." This is illustrated, according to him, by connectionist models that seem to show that "such processing is reliable, but that its reliability is not entirely a function of operations on person-level representational contents, such as evidence and inference rules are understood to involve." If this is correct, "the reliability of such processing cannot be understood in evidentialist terms, such as evidential 'fit' and 'support.'" The reliability of our cognitive systems, then, does not seem to depend on facts about evidence and, evidentialism, as a result, is undermined by contemporary cognitive science.

4.4. The operations of the intuitive, non-inferential, and instantaneous natural cognition

Robert McCauley similarly proposes that the description of the operations of the human mind that has been revealed to us by cognitive science shows that evidentialism is an untenable theory of justification and knowledge. He illustrates this with the scene, from Anthony Trollope's novel Barchester Towers, of an exchange between Mr. Quiverful and Mr. Slope in which the former grasps instantaneously, without reflection or inference, the latter's intention to dissuade him of his plans to accept a position which had been offered to him by Mr. Slope himself. Mr. Quiverful, Trollope tells us,

> saw at a glance that his brilliant hopes were to be dashed to the ground and that his visitor was now there for the purpose of

> unsaying what on his former visit he had said. There was something in the tone of voice, something in the glance of the eye, which told the tale. Mr. Quiverful knew it all at once.[61]

This exchange illustrates, according to McCauley, the operations of natural cognition, which is intuitive, non-inferential, instantaneous. Its operations are to be contrasted with those of unnatural cognition, characteristically reflective and inferential. McCauley associates unnatural cognition with knowledge on the basis of evidence and natural cognition with knowledge that doesn't arise from reflection and "carefully weighing the evidence" (13), but from a capacity to grasp things (even complex ones) instantaneously, "in a flash, as when lightning suddenly illuminates everything around us at night" (12). Knowledge from natural cognition does not arise from appropriate response to the evidence. It arises from "intuitions" that are "far more elaborate and refined than the readily accessible evidence supports." Intuitions, though they may amount to knowledge, as in Mr. Quiverful's case, are "insufficiently supported by the available evidence". But, more than that, claims McCauley, "often it is not obvious what evidence suggested them" (14).

4.5. Ad Barrett, Greco, and McCauley

So Barrett's, Greco's, and McCauley's objections to evidentialism from cognitive science have similarities, but are not identical. It seems that Justin Barrett's core claim against evidentialism is:

> (JB) Evidence is always filtered and distorted by the operation of mental tools. We never have direct access to evidence but only processed evidence—memories.

Perhaps Barrett's objection can be put more schematically as follows:

(1) If evidentialism is true, evidence cannot be filtered and distorted by our cognitive system.

(2) Evidence is distorted by our cognitive system.

(3) Evidentialism is false.

[61] *Apud* McCauley, 12.

John Greco's core claim against evidentialism seems to be found in the following passage:

> (JG) Recent empirical studies make it doubtful that paradigm cases of knowledge,[62] such as perceptual knowledge, memory knowledge, and inductive knowledge, can be understood entirely in terms of person-level representational states, as evidence is understood to be.

Perhaps Greco's objection can be constructed as:

(1) If evidentialism is true, knowledge can be understood entirely in terms of person-level representational states.

(2) Knowledge cannot be understood entirely in terms of person-level representational states.

(3) Evidentialism is false.

And Robert McCauley's objection seems to be that

> (RM) Knowledge through natural cognition results from a response to our environment that goes beyond what the evidence available to us indicates.

Perhaps McCauley's objection can be expressed more formally as follows:

(1) If evidentialism is true, knowledge through natural cognition cannot go beyond what the evidence available to us indicates.

(2) Knowledge through natural cognition can go beyond what the evidence available to us indicates.

(3) Evidentialism is false.

All these authors are pointing to cases of knowledge without evidence. Greco points to cases of blindsight, among others, and seems to be claiming that the sort of input involved in knowledge does not involve evidence as this notion is normally understood: if knowledge

[62] As noted in chapter 2, evidentialism is primarily a theory about epistemic justification and, only secondarily, a theory about knowledge. Still, most evidentialist are likely to embrace the view that knowledge requires justification. As a result, they will want to reject any view that entails that knowledge does not require evidence.

depends on sub-personal inputs, evidentialism cannot be true, for these inputs do not indicate the truth of *p*, which is something evidence is expected to do. McCauley's claim is similar. He seems to be pointing to cases of knowledge without evidence and the reason why he believes belief-formation in the case of knowledge he discusses does not occur in response to evidence is because he can't point to something in that process that would resemble what he understands as evidence. And Barrett's objection also seems to rest on the idea that we know things without evidence, for even if the initial input involved in cases of knowledge can be properly said to consist of evidence, this initial input goes through a process of filtering and modification that desconfigures the original input and, thus, it does not seem that we can say that the belief is formed in response to the original information (the evidence). In the end, then, it seems that Greco's, McCauley's, and Barrett's claims are similar. Theirs is a claim about the nature of evidence. They conceive evidence as understood by the evidentialist as having a certain nature. But they find cases of knowledge that do not involve evidence of that sort.

Thus, in light of (JB), (JG), and (RM), and the corresponding arguments, what can be said in defense of evidentialism, if anything? It seems that the discussion of section two of this chapter and our review of the literature on the nature of the evidence and on evidentialism in chapters one and two provide us with the resources that we need to articulate possible evidentialist responses to these objections. First and foremost, all three objections can plausibly be said to suffer from an excessively restrictive conception of evidence. But, as we saw above and in the previous chapters:

- Most evidentialists today hold to the view that evidence consists in mental states or the propositional content of experiences;
- As Conee and Feldman put it, "evidentialists need not appeal only to conscious occurrences as justification" (2011, 304);
- The problem of stored beliefs has led many evidentialists to emphasize that dispositions constitute evidence and to embrace a dispositionalist account of belief and evidential support;

- Evidentialists need not appeal only to dispositions that constituted at some point information learned through a conscious process, *i.e.*, that entered the mental life of the subject reflectively before they were stored;[63]

- Evidentialists need not appeal only to information that was stored exactly as they entered the subject's mental life and that didn't go through any process of filtering and modification.

To recapitulate Matthew Frise's response to the problem of stored beliefs (2017a, 2017b, 2018), the standard picture of our memory that has emerged from research in cognitive psychology goes counter to the common view that memory is like a storehouse where beliefs are kept unchanged until they are recalled by the subject. Rather, after information enters our memory, it goes through three stages of processing that will significantly alter the initial input (2018). First, the information is *encoded*, *i.e.*, a process of content alteration of the information begins, with the "gist" of the information being extracted and stored. Second, there is *consolidation*, *i.e.*, content alteration continues, but now with the generation of additional content (about general features of the information). Third, the information that went through the process of modification in the two first satge, with elimination of peripheral content in the first, and addition of new content, in the second, is now *retrieved*. The process of retrieval, however, is also far from a passive one. In this stage, the transformation of the stored information into beliefs is constrained by the context, adding to the process that makes what is recalled only partially resemble what was initially stored. This context has three components: (a) it involves the cues that prompt recollection, leading to variation in both the quantity and quality of the retrieved information according to the manner in which the information is cued; (b) the information previously retrieved can affect the quantity and quality of the information extracted in the next round of retrieval; and (c) Through the process known as source monitoring, we are able to distinguish between

[63] According to Frise: "there is reason to deny that a stored belief must at some time have been occurrent. Information that enters memory is normally altered in at least three stages of memory processing. If memory stores nearly as many beliefs as we think it does, we should allow that we believe the content resulting from this processing. But much of this content has never been occurrently endorsed. So a stored belief need not have been occurrent at any time" (2017a, 491).

the retrieved information originated from what we experienced or from what we merely imagined.

Throughout this process, from the beginning of stage one until the completion of retrival, information is altered, with the final result looking a lot different from what entered one's memory at the beginning of the first stage. Yet, nothing in this process undermines the reliability of memory. While differing in important ways from the original information, the belief formed at retrival tends to be true. Source monitoring plays a major role in securing the reliability of memory. While differing in important ways from the original information, the belief formed at retrival tends to be true. Source monitoring plays a major role in securing the reliability of memory by indicating to us whether the origins of the retrived information are to be found in our experience or our imagination.

After presenting this picture of memory, Frise claims that it falsifies representationism about belief, with its reliance on a picture of memory that resembles a storehouse. On the representationalist picture, beliefs – *i.e.*, mental representations – are stored until they are retrived. But, if, as cognitive psychology seems to show us, beliefs are not simply stored in memory – they are "dismantled," reconfigured, and reassembled, emerging from recollection bearing little resemblance with the original belief – than beliefs cannot be representations. And, if they are representations, there are no stored beliefs. As a result, Frise searches for an alternative account of the nature of belief that can give us the result that there are stored beliefs even though the information that enters our memory and forms the content of our beliefs undego the process of modification described above. He finds the compatibility of these two desiderata in dispositionalism. On the dispositionalist account of belief, a belief is a set of dispositions, or, more precisely, "a suitable combination of dispositions pertaining to *p*" (Ibid., 69). So, it doesn't matter, on this picture, whether mental representations are filtered and dismantled. What matters is whether we will have the suitable dispositions pertaining to p. Frise claims that, either way, the problem of stored beliefs for evidentialism is solved: if beliefs are mental representations, then there are no stored beliefs, and, hence, no problem of stored beliefs; if beliefs are dispositions, then we do have stored beliefs, and they are formed by the set of dispositions one has.

Recall that the problem of stored beliefs is, roughly, one of identification of the source of evidence for these beliefs that could satisfy the evidentialist conditions for justification. On Frise's suggested dispositionalist condition on justification, a stored belief that *p* is justified when "*S* has a suitable combination of dispositions pertaining to *p*" (70). To use one of Frise' examples: when your cellphone rings and your automatic response is to reach into your pocket to pick it up, you weren't thinking about the location of your cellphone. You non-occurrently believed your cellphone was in your pocket because you had the suitable combination of dispositions. And that belief was justified. Even before your cellphone rang, you were disposed to have a recollecting experience of the location of your cellphone and feel confident about it. Hence, on this picture, even when one is not thinking about *p*, one may be justified in her belief that *p*, provided she has the right set of dispositions pertaining to *p*. And the fact that the information stored in memory is filtered and distorted, as Justin Barrett rightly claims, doesn't do anything to discredit evidentialism. For what constitutes the evidence that justifies one's belief is not the initial information, but the set of dispositions one has when forming the belief. As Barrett will be the first to aknowledge, our cognitive faculties are reliable. Our beliefs are appropriate responses to the world out there. While the initial bit of information (the information that evidences how things really are out there) may go through a process of modification, whatever remains of it (with the pairing and additions of content it undergoes in our mind) will, in the end, be reliable indicators of how things really are.

And when Greco says that evidence is understood by evidentialists to be person-level representational states, it is not clear which philosophers he has in mind as proponents of this view, but it doesn't seem to capture the views of some of the main proponents of evidentialism today, such as Byerly, Feldman, Frise, and McCain. All of these epistemologists view dispositions as evidence, and argue that justified beliefs can be formed on the basis of such evidence.

* * *

We are now in position to see where the three arguments fail to establish their conclusions.

Recall that the argument underlying (JB) is:

(1) If evidentialism is true, evidence cannot be filtered and distorted by our cognitive system.

(2) Evidence is distorted by our cognitive system.

(3) Evidentialism is false.

The problematic premise here is (1) (in fact, (2) seems true and evidentialists like Frise studying the science of memory acknowledge that). As Frise has shown, the distortion of the original information need not pose a problem for evidentialism. For evidentialists can plausibly conceive dispositions as evidence and claim that beliefs consist in sets of dispositions.

Recall now the argument underlying (JG):

(1) If evidentialism is true, knowledge can be understood entirely in terms of person-level representational states.

(2) Knowledge cannot be understood entirely in terms of person-level representational states.

(3) Evidentialism is false.

Here, too, evidentialists will grant (2). But they will find premise (1) problematic. The basis for this is that the favored solution to the problem of stored beliefs has led evidentialists to emphasize that sub-personal states can be evidence. Moreover, cognitive psychologists working on memory have presented a picture of the functioning of the human memory in which that seems to go against the representationalist view of belief. So evidentialists will probably want to say here that (1) does not reflect their understanding of evidence, which can be in fact understood as including sub-personal states in the form of dispositions.

And, finally, according to the argument unpacked from (RM):

(1) If evidentialism is true, knowledge through natural cognition cannot go beyond what the evidence available to us indicates.

(2) There are cases of knowledge through natural cognition that go beyond what the evidence available to us indicates.

(3) Evidentialism is false.

Here the evidentialist can grant (1), but not (2). Again, although evidentialism is primarily a theory about justification, most evidentialists are unlikely to concede the possibility of there being unjustified knowledge.[64] And the reason why the evidentialist will reject (2) is because she will claim that the cases of knowledge through natural cognition presented by McCauley are not really cases in which one is not obviously failing to meet whatever specific conditions the evidentialist defends for propositional and doxastic justification. For McCauley seems to have excessively restrictive view about the nature of the evidence and of evidential support. An evidentialist can say that Mr. Quiverful had certain dispositions or background beliefs about human behavior that, in conjunction with Mr. Slope's behavior, demeanor, and the circumstances of his visit, led him to see (without any inference) that "his brilliant hopes were to be dashed to the ground." Perhaps (as suggested by McCain) by having a disposition to have the seeming that his experience of Slope's behavior, demeanor, and the circumstances of his visit, was best explained by his intention to "[unsay] what on his former visit he had said." Or perhaps Mr. Quiverful was (as suggested by McAlister and Dougherty) *perceiving* tacit evidential relations that made him see that his plans had failed. Or perhaps (as suggested by Byerly) Mr. Quiverful was simply disposed to believe that his hopes were lost in light of his total evidence that included dispositions or background beliefs about human behavior in conjunction with Mr. Slope's behavior, demeanor, and the circumstances of his visit.

4. Conclusion: Toward an Improved Model of the *Sensus Divinitatis*

In this chapter, we attempted to bring together the discussions of the previous chapters on evidence, evidentialism, and cognitive science of religion in order to evaluate the compatibility (or lack thereof) of religious evidentialism with what cognitive scientists have been saying about the origins of religious beliefs and about the functioning of the human mind more generally. We began with a survey of the literature on reformed epistemology and on the *sensus divinitatis*. We then moved to a discussion of the various ways

[64] An exception here might be knowledge-firsters (see chapter 1) who consider themselves evidentialists about justification but who endorse the view advocated by Lasonen-Aarnio (2010) that there can be unreasonable knowledge.

evidentialists can understand the nature of evidence and evidential support (building on what was discussed in chapters two and three) in order to prepare the ground for an examination of the three CSR approaches discussed in the previous chapter in light of the suggestion that their findings may be incompatible with what evidentialists can say about the justification of religious beliefs. The result of this examination seems to be clear: there is nothing about those findings that preclude an evidentialist understanding of how certain religious beliefs can be justified.

But perhaps the problem lies deeper. Perhaps the problem for evidentialism lies on what cognitive science shows us about how information is processed by our cognitive equipment. Perhaps, as Barrett, Greco, and McCauley have suggested, the picture that cognitive science has provided us with respect to how we form beliefs that amount to knowledge shows us that our beliefs in those circumstances do not involve what evidentialists could plausibly call "evidence." On the basis of the current literature on the nature of evidence and on the attempt made by evidentialists to deal with problems such as how to account for stored beliefs, we showed that this claim rests on unnecessarily limited notions of the nature of evidence and of evidential support.

Finally, it seems that our survey of the literature on the nature of evidence and on evidentialism, on the one hand, and of the three CSR accounts, on the other, has revealed to us some limitations of a recent attempt to reform reformed epistemology. While reductivism can account very well for the findings of the first CSR approach, and while it doesn't seem to be incompatible with the CSR accounts that we examined, it doesn't seem to account as well for the findings of the dispositional and preparadness CSR approaches as other potential versions of evidentialism that make dispositions central to the way the nature of evidence and evidential support are conceived. Thus, while the reductive model, which relies on the mechanism of tacit perception of evidential relations, can account for the findings of the CSR attribution account quite well, it cannot accommodate the findings of the other two approaches so well – or at least not as well as potential alternative evidentialist theories. A successful model with respect to all CSR approaches should make dispositions central in how the nature of evidence and evidential support are understood. And, in fact, the dispositionalism of Byerly (and perhaps Frise) and, to a lesser extent, the inclinations view of

seemings and the explanationism of McCain seem to deliver results that are more consistent with the findings from CSR and can thus, potentially, provide better grounds for the construction of a successful model of the *sensus divinitatis* that take the findings of all major CSR approaches to the origins of religious belief as important factors in the evaluation of the success of such a model.[65] Let me conclude by saying something about the direction the development of such models could take in these alternative evidentialist approaches.

The dispositional approach developed by Byerly (and perhaps Frise) seems to be able to account very successfully for all the CSR approaches. It seems therefore to be an excellent candidate for a successful model of the *sensus divinitatis*. On this model, God designed us in such a way that, as we develop cognitively, we would become disposed to believe certain religious propositions in light of our total evidence. In terms of the attribution approach, we would become disposed to believe that God loves us in light of our total evidence that includes the visual sensation of a sunset. In terms of the dispositional CSR approach, we would simply become more receptive (it would fit well with our naturally developing set of dispositions) to accept certain religious ideas and believe them in light of our total evidence. In terms of the preparadness approach, as we develop cognitively, we would become disposed to believe certain teleological and religious propositions in light of our total evidence. A potential problem for this model is that, if we take beliefs to be dispositions (something one might be motivated to do in order to avoid the problem of stored beliefs), then Byerly's model seems to reduce to doxastic conservatism: one is justified in believing *p* because one already believes *p*.[66]

[65] I am not suggesting that a successful model of the *sensus divinitatis must* accommodate all the findings from CSR. For one may have theological and philosophical considerations in mind that may override some of these findings. After all, CSR is a nascent discipline. It has provided extraordinary insight into a number of important questions explored by religious epistemologists, but, as it is part of the very nature of the scientific endeavor, its findings (or at least a large portion of them) are provisional and potentially falsifiable (which is not to say that philosophical and theological data are not, though philosophical and, primarily, theological data tend to rely on certain findings or presuppositions that are by their own nature less likely to be overridden by new scientific findings that may justifiably be regarded as provisional and contingent).

[66] Byerly (2014, 417) adds to his dispositionalist account of justification the condition that one should be "strongly" disposed to believe, rather than simply be disposed to believe, in order to differentiate his theory from epistemic conservatism. It is not clear to me, however, that this addition would eliminate the problem just mentioned if we take beliefs to be dispositions.

An explanationist account along the lines developed by McCain could be used very successfully to model a *sensus divinitatis* that can accommodate the findings of the preparadness approach. We would naturally develop a set of dispositions to have seemings that certain teleological and religious propositions are the best explanation for the mental states we have. When confronted with questions such as why mountains, stars, or people exist, we would be naturally disposed to have the relevant kinds of religious seemings.

The explanationist account also seems capable of delivering the correct results with respect to the attribution approach (we may be disposed to have seemings that the supernatural agents that we believe we are detecting are the best explanation for our experiences). It is unlikely that explanationism would be so successful in accommodating the dispositional CSR account, however, but it could accommodate it nonetheless (perhaps some of the religious beliefs described by that account would be accompanied by dispositions to have seemings that those religious propositions are the best explanation for the relevant mental states that accompany those beliefs). But suppose explanationism would have significant problems accommodating the dispositional CSR account. The explanationist could perhaps claim that beliefs produced exclusively via the mechanisms described by the dispositional CSR account do not produce justified beliefs in and of itself, which would not be an implausible claim. As a result, one could perhaps argue that an explanationist account of the *sensus divinitatis* could actually deliver the right results with respect to the models that are more likely to produce justified religious beliefs.

Third, there is the inclinations view of seemings, according to which the "feeling of truth" or the "forcefulness" that accompany seemings should not be understood as resulting from a belief one holds, or from an experience with propositional content, but from an inclination to believe. I distinguished above dispositions from inclinations. The former are non-occurrent mental states one is not aware of; the latter, on the other hand, are occurrent mental states one is aware of. An account of the *sensus divinitatis* in terms of inclinations would have to presuppose that the inclinations that produce religious beliefs follow directly from the dispositions that constitute the mechanism (the *sensus divinitatis*) designed by God to give us justified religious beliefs. So, this model may not add anything relevantly different from the previous, dispositional, model. It might, nonetheless, be immune to potential

problems with other phenomenal conservative accounts of the *sensus divinitatis* (in terms of beliefs or experiences, rather than inclinations) that seek to accommodate the findings of CSR.

The problem with phenomenal conservative accounts of the *sensus divinitatis*, or at least of an account that seeks to be consistent with the findings of the three CSR accounts, is that the process of filtering and distortion of the information that enters our cognitive equipment raises the specter of something similar to the cognitive penetration of perceptual beliefs, *i.e.*, the idea that our perception can suffer the influence of beliefs (see Tucker, 2013, for discussion of the problem and possible solutions). Analagously, if all our reflective beliefs are, according to the model of belief formation described by Barrett above, subjected to the influence of non-reflective beliefs, this might mean that all reflective beliefs are subjected to something akin to cognitive penetration. This problem would not afflict dispositional accounts of the *sensus divinitatis*, since, on those models, dispositions are evidence and beliefs can be construed as dispositions, rather than as representations. Something similar can perhaps be said with respect to explanationism.

In any case, these are just some suggestions of directions in which the pursuit of an adequate model of the *sensus divinitatis* can go. More details need to be filled for each of these alternative models before they can be shown to be superior to the alternative models developed by Plantinga, Tucker, and McAllister and Dougherty. And that's a task for another occasion.

PART II

RATIONALITY AND DEFEATERS

RATIONALITY

The concept of rationality can be applied to persons, actions, desires, and so on. The kind of rationality that concerns us in this essay is rationality applied to beliefs. This is normally called "epistemic or theoretical rationality" – rationality about the appropriateness of accepting a certain belief.

Epistemologists have traditionally spoken of rationality and justification interchangeably. Thus, on this view, unjustified or irrational beliefs may be taken to be beliefs that fulfill the conditions for epistemic justification. There are different such accounts. A belief may be said to be unjustified – and hence irrational – because it doesn't fit the evidence one has (evidentialist condition), or because it is not reliably formed (reliabilist condition), or not causally connected with the object (causal condition), or not produced by a faculty functioning properly according to a design plan (proper functionalist condition), or not produced by an exercise of one's intellectual virtues (virtue condition), and so on.

It is important to note, however, that not all epistemologists see rationality and justification as being interchangeable notions. Robert Audi (2011), for instance, sees the rationality of a belief as depending on internal coherence and good grounding on one's experiences. Justification, on the other hand, is, on his view, a notion related to achievement, to the proper response to one's experiences. And Audi distinguishes between these and a third category, which is often taken to be closely related to rationality, but which is stronger: reasonableness. While rationality requires internal coherence and consistence with one's experiences (2004, 33), reasonableness requires coherence within one's belief system and with one's experience and that, in forming the belief, the subject manifests intellectual virtues such as perceptiveness, good judgment, and significant capacity for good reasoning (2004, 40). Another way of formulating the distinction is that rationality requires that one's beliefs be consonant with reason, whereas reasonableness requires that one's beliefs be supported and governed by reason (2011). Audi summarizes his views on how the concepts of rationality, justification, and reasonableness are related as follows:

> A natural and promising way to begin to understand rationality is to view it in relation to its sources. The very same sources yield justification, which is closely related to rationality. These sources are also central for reasonableness, which implies rationality but is a stronger notion. Our reasonable beliefs, like our justified ones, are rational, but a belief that is rational—at least in the minimal sense that it is not irrational—may be (beyond avoiding inconsistency and other clear defects) simply plausible to one, sometimes in the way a sheer speculation often is, and may fail to be justified or reasonable, as one may later admit (2004, 18).

Another important distinction is between internal and external rationality (Plantinga, 2000, 110-112, Bergmann, 2009). Internal rationality involves the formation of beliefs in response to experience. A belief is internally rational when it is an appropriate response to one's mental states. External rationality depends, on the other hand, on what happens causally prior to experience. A belief is externally rational when it is formed according to the proper function of the cognitive faculties involved. Perhaps an example will help to make more vivid where the difference between these two types of rationality lies. As an example of an internally rational and externally irrational belief, Plantinga cites Descartes observation that there are "people whose cerebella are so troubled and clouded by the violent vapours of black bile, that they . . . imagine that they have an earthenware head or are nothing but pumpkins or are made of glass" (2000, 111). Provided that the experience of those people supported in a coherent way (in light of their total evidence) the belief that their heads were made of glass, they were internally rational in holding such a belief. But their beliefs were externally irrational, for the cognitive faculties involved in the formation of such a belief were malfunctioning. Properly functioning cognitive faculties do not deliver experiences to the effect that us humans have heads made of glass. Thus, one may be internally rational while being externally irrational, and vice-versa: one can be internally irrational while being externally rational – *i.e.*, one can fail to properly respond to the evidence available to her while her belief is formed by properly functioning cognitive faculties.

While I find Audi's distinction between the notions of rationality, justification, and reasonableness an important step in the right direction, I am using "rationality" and "justification" interchangeably and have been avoiding using the notion of reasonableness. The reason for this is that, while I do think there are important distinctions between these notions, I am not convinced Audi's characterization of them and explanation of how they are related fully capture what they are and how they are related, though, again, I believe major progress has been made by the way he explicates those notions. For this reason, I have been using "rationality and justification" interchangeably in this essay.

DEFEATERS

"Knowledge," "warrant," "justification," "rationality," and other properties that denote positive epistemic status are defeasible. This means that the positive epistemic status of beliefs can be eliminated or reduced given the actualization of certain conditions. Whatever (whether information, evidence, reason, truth, etc.) actualizes the condition of loss or reduction of positive epistemic status of beliefs can be called a "defeater."[67]

The term "defeasibility" has its origins in the legal literature. Contracts, for example, can be canceled, invalidated, or defeated given the actualization of certain conditions. In philosophy, the word "defeasibility" was initially employed in ethics to denote the possibility of suspension or elimination of certain moral obligations given the presence of other factors with the power to override these *prima facie* valid obligations.[68] This terminology became widely used in epistemology after Edmund Gettier (1963) presented his counterexamples to the traditional analysis of knowledge. One of the first proposed solutions to the Gettier problem involved the formulation of an analysis of knowledge in terms of undefeated justified true belief. The theory developed in this attempt to provide a definition of knowledge in terms of absence of defeaters became known as the "defeasibility theory."

[67] This definition of defeater is found in Michael Sudduth (2008). It seems to capture what is essential to most types of defeaters, and the various attempts in the literature to define epistemic defeat seem to be subsumed under that definition.

[68] For a brief presentation of the legal and ethical context in which the notion of defeasibility emerged, see Sudduth (2008).

The defeaters discussed by proponents of the theory of defeasibility are propositional defeaters, which are external to the subject's mental life, but which, if they were believed by her, would defeat her justification to believe in the proposition in question.[69] For defeasibilists, justification that is good for knowledge is justification that is resistant to the addition of truths to the subject's mental life. If there are truths outside the subject's belief system that, once believed, would eliminate their justification to believe in the target proposition, the subject lacks knowledge. Most of the literature on defeaters today, however, is concerned with defeaters that are part of the mental life of the subject, and, for this reason, are called "psychological defeaters" (Lackey, 2008), "mental states defeaters" (Bergmann, 2007), "internal defeaters" (Sudduth, 2008), "evidential defeaters" (Lyons, 2011) or, simply, "defeaters" or "overriders" (Klein, 1981). A third type of defeater, which has been, as we shall see below, the subject of growing interest in the literature, is that of normative defeaters.

Internal defeaters (hereafter simply defeaters) come in many varieties. John Pollock (e.g., 1986) is the philosopher responsible for the initial (and still widely accepted) classification of defeaters as rebutting defeaters and undercutting defeaters. The first are reasons that the agent has to regard one of her beliefs, *c*, as false and thus give up that belief. For example: Carlos sees a sheep in the pasture and forms the belief that there is a sheep in the pasture. But the owner of the pasture, a trustworthy man, tells Carlos that there is no sheep there, but rather an English shepherd dog that at a distance looks like a sheep (Plantinga, 2011: 164). The owner's testimony gives Carlos reason for believing that his belief that there is a sheep in the pasture is false. Undercutting defeaters, on the other hand, instead of being reasons that lead us to believe that *c* is false, are reasons that lead us to believe that we do not have good reasons to believe that *c*. For example: Márcio sees someone leaving the house across the street and forms the belief that Paul is leaving home. But Márcio hears that Peter, Paul's twin brother, arrived at Paul's house last night. This information is not sufficient

[69] This subjunctive formulation is avoided by some epistemologists (such as Klein, 1981) with the aim of escaping from the conditional fallacy (Shope, Robert (1978), The Conditional Fallacy in Contemporary Philosophy, 75 (8): 397-413). I am grateful to Rodrigo Borges for drawing my attention to the controversy surrounding the formulation of the defeasibility clause in conditional terms.

for Márcio to acquire a rebutting defeater for his belief that Paul was leaving home. However, it is a reason to make Márcio doubt that seeing someone like Paul leaving the house is a good enough reason for him to believe that Paul was leaving home, thus compelling Márcio to suspend judgment with respect to whether Paul was leaving home (Plantinga, 2011: 165).

Alvin Plantinga is the epistemologist who has given the most significant contribution to the expansion of our vocabulary about defeaters. Plantinga accepts Pollock's distinction between rebutting defeaters and undercutting defeaters, but uses the following additional types of defeaters in his epistemological work (*e.g.*, 1994, 2000, 2011): warrant defeaters, proper-function-rationality defeaters, Humean defeaters, purely alethic rationality defeaters, potential defeaters, defeater-deflectors, defeater-defeaters, neutralizing defeater-defeaters, intrinsic defeater-defeaters, extrinsic defeater-defeaters, partial defeaters, and “optimistic overriders.”

Warrant, according to Plantinga's (2000) definition, is the quality that distinguishes knowledge from mere true belief. *Warrant defeaters*, therefore, are defeaters that eliminate this quality, which means that there can be defeat of warrant without there being defeat of rationality. Plantinga (2011) cites the well-known barn case as a case in which Henry's knowledge that he is seeing a barn is defeated by an element of the environment (the fake barns that surround the true barn), not by a belief or experience. *Rationality defeaters*, on the other hand, operate through beliefs or experiences and would, in Plantinga's (1994) view, come in two varieties: *proper-function-rationality defeaters* and *purely alethic rationality defeaters*. A belief *d* is a defeater of the first type of *c* if the proper functioning of our cognitive faculties requires that we give up *c* when we acquire *d*. The need to establish a category of purely alethic rationality defeaters comes from the fact that the proper functioning of our cognitive faculties may require that a certain belief be formed on the basis of wish-fulfillment rather than pursuit of truth. An example is the formation of a belief, by a terminally ill patient in circumstances where all the evidence points to the impossibility of her recovery, that he will recover. And in situations where the proper functioning of our cognitive faculties leads us to form optimistic beliefs to the detriment of what the evidence indicates, that is an instance, according to Plantinga (2000), of "optimistic overrider."

Humean defeaters are defeaters of rationality that produce global skepticism (2002). A subject, for example, who believed that the pill he had just taken had the effect of producing complete loss of reliability of the cognitive faculties of those who ingested it would be rationally compelled to suspend judgment about the reliability of her cognitive faculties. In such a situation, the subject is said to have acquired a Humean defeater. Plantinga's Evolutionary Argument against Naturalism (2011) is precisely an argument that aims to show that the naturalist acquires a Humean defeater by noting the low probability (Pr) that our cognitive faculties are reliable (R) given that we are the product of an evolutionary process (E) not guided by God (N) – that is, Pr (R / E & N) is low. Humean defeaters are defeaters of global reliability. There are, however, reliability defeaters[70] that target more specific faculties or domains of our cognitive life, producing local, rather than global, skepticism.

Defeaters are relative to noetic structures, the rest of what we believe. Thus, what constitutes epistemic defeat for one subject may not constitute defeat for another subject in the same context. In the sheep case, the owner's testimony that there is no sheep in the pasture, but rather a dog that, at a distance, resembles a sheep, is a defeater because Carlos relies on the owner's testimony. But epistemic defeat would not materialize for another person who possessed a belief that the owner is an irreverent person who likes to make people believe falsely that there is no sheep in his pasture.

There are situations in which the subject believes the defeater but does not see the connection between the defeater and the defeatee. In these cases, the subject is in possession of a *potential defeater*, a dormant defeater that does not produce its defeating effect until the subject sees the connection (Plantinga, 2000: 361). When Frege read the letter that Bertrand Russell had written to him, alerting him of the implications of his belief that "for every property *p* there is a set of things that have *p*," which was up to that point a potential defeater (the implication that that set would exemplify and not exemplify itself at the same time) became an effective defeater for the belief in the existence of that set.

[70] The notion of reliability defeaters is employed by William Talbott. See, for example: A Non-Probabilistic Principle of Higher-Reasoning. Synthesis, 193: 3099-3145, 2016.

Defeaters that defeat other defeaters are called defeater-defeaters. As in the case of defeaters, there are several types of defeater-defeaters. One of them is what Plantinga calls *neutralizing defeater-defeaters* (1994). These are defeater-defeaters that, instead of defeating the original defeater, produce a neutralizing effect of their defeat potential, so that both the original belief, the defeater, and the defeater-defeater coexist in the belief system of the subject.[71] Defeater-defeaters can be *intrinsic* or *extrinsic* (1994). Defeater-defeaters of the first type are beliefs that have a greater degree of warrant than potential defeaters. Plantinga illustrates the way in which this kind of defeater works with the case of a subject charged with a crime who finds himself with all external evidence pointing to his guilt but who vividly remembers not having been at the crime scene at the time it occurred. The memory of the subject in this case constitutes an intrinsic defeater-defeater (2000: 371). Extrinsic defeater-defeaters, on the other hand, are the defeater-defeaters that lack this property that characterizes the intrinsic defeaters.

Defeater deflectors are beliefs that are already part of the belief system and that, as long as they are part of that noetic structure, it is not possible to form rational belief in *c*. In the case of the sheep, the owner's testimony to the effect that there is no sheep in the pasture, but rather a dog that, at a distance, resembles a sheep, is a defeater for Carlos's belief that there is a sheep in the pasture. But let us suppose that, before Carlos' conversation with the owner, the owner's wife tells Carlos that her husband has a habit of telling everyone who passes by that there is no sheep in the pasture. Since Carlos believes she is telling the truth, the testimony of the owner does not become a defeater. The testimony of the owner's wife functions as a deflector of the defeater. The difference between defeater-defeaters and defeater deflectors, therefore, is that while the former presupposes that one has a defeater *d* for the belief *c*, the latter prevents *d* from becoming a defeater in the first place.

Finally, *partial defeaters*, as the name suggests, are defeaters that do not require the suspension of judgment with respect to a particular belief, but rather reduction of the conviction with which it is held (Plantinga, 2000: 362, note 3). Moreover, in Plantinga's epistemological work we also find the original attempt to elaborate principles of epistemic

[71] For examples that make the idea clearer, see: (1994) Naturalism Defeated, pp. 36-7.

defeat. Plantinga (1994) has left this work unfinished, now it being incumbent on the present and future generations to continue the work of elucidating the principles that govern our mental life with respect to epistemic defeat.

Pollock's and Plantinga's work on epistemic defeat was supplemented by that of Michael Bergmann. An initial distinction made by Bergmann (2007) is between *newly acquired* defeaters and *continuing* defeaters. The first type consists of defeaters acquired at *t* that make justified beliefs before *t* no longer justified at *t*. The second type concerns newly acquired defeaters that continue to produce the original defeating effect. Newly acquired defeaters, on the other hand, may be divided into newly acquired *state* defeaters or newly acquired *power* defeaters. The former, which are the most common types of defeaters discussed in the literature, are newly acquired mental states, while the latter are mental states that are already part of the subject's belief system but which only subsequently acquire power to function as defeaters. There is also what Bergmann (2007) calls no-reason defeaters, which defeat only beliefs not based on reasons or evidence. The no-reason defeat occurs when the subject realizes that she does not have reasons for a certain belief, although she should have them. This type of defeater is characterized, according to Bergmann, as a particular case of defeater that cannot be subsumed under the categories of rebutting defeaters or undercutting defeaters.

More recently, some epistemologists (such as Jennifer Lackey (2008), Sanford Goldberg (2015), and Jack Lyons (2011)) have argued for the existence of what they call "normative defeaters." Internal defeaters, as we have seen, are beliefs or experiences (or perhaps even propositional attitudes such as suspension of judgment, as suggested by Bergmann (2007)) that would negatively affect the justification, rationality or reasonableness of holding other beliefs. Normative defeaters, on the other hand, consist, in Goldberg's characterization (2015), in the violation of legitimate expectations given the social role the agent plays in a particular context. Or, according to Lyons, "a normative defeater is a factor that disrupts justification in cases where an agent believes something only because she hasn't done something she should have done" (Lyons, 2011: 307, note 6). The cases that Goldberg and Lyons have in mind involve situations such as, for example, a subject who fails to verify her voice mail or to read her correspondence. Normative defeat would occur in such cases if the

content of the message or of the letter brings evidence that, if believed by the agent, it would defeat the positive epistemic status of a particular belief. For Goldberg, people are expected to check their voice messages and read their letters, and evidence contained in these media could have a negative effect on the epistemic status of beliefs, even if they are not part of our mental life. Violation of these social norms may keep the agent from obtaining defeating evidence in the sense of internal defeat, but it does not prevent her, according to these authors, from acquiring a normative defeater. J. Adam Carter (2015: 13) illustrates this with the following case: a father decides to take the family to the zoo and, as they approach the zoo, he fails to read signs to the effect that the zoo will be closed that day. According to these authors, the justification of the father's belief that he would spend the day at the zoo with his family is defeated by the signs, even though he has not read them. The father, on this view of defeaters, should have paid attention to the signs. By failing to do so, he acquires a normative defeat.

In addition to the considerations listed above about defeaters, other relevant questions might be mentioned here in passing: Can all types of defeaters discussed above be subsumed under the categories of rebutting and undercutting defeaters? Can unjustified beliefs defeat? Can group beliefs be defeated? With respect to the first question, there are attempts in the literature to defend the existence of hybrid accounts of defeaters (*e.g.*, Kotzen, 2010). The answer to the second question seems to be "yes" if defeat is conceived in purely psychological terms, and "no" if conceived in epistemic terms.[72] And recently criticisms have been raised against the prevailing conceptions of collective epistemology on the basis of their apparent inability to accommodate defeaters of group knowledge (Carter, 2015).

Finally, it is important to mention that criticisms of the traditional notion of epistemic defeasibility have emerged in recent years. According to Maria Lasonen-Aarnio (2010), externalist theories of knowledge have illegitimately incorporated the internalist apparatus of epistemic defeat. In particular, theories of knowledge that incorporate a safety condition to the conditions of truth and belief for knowledge would not be able to explain how safety,

[72] I thank Rodrigo Borges for suggesting this distinction between epistemic (loss of belief in response to evidence) and psychological (loss of belief for non-epistemic reasons).

and therefore knowledge, could be lost in cases of misleading evidence. The conclusion defended by Lasonen-Aarnio is that, in such cases, although reasonableness is lost, knowledge is preserved. Max Baker-Hitch and Matthew Benton (2015) seek to complement Lasonen-Aarnio's criticism of the notion of knowledge defeat by posing problems for the incorporation of the notion of defeat into other externalist approaches and arguing that internalists, too, have serious difficulties in formulating a plausible account of epistemic defeat since, on a Bayesian formalization of undercutting defeat, there are difficulties with the idea that a defeater can "lower one's probability for some *p*, but for the evidence (apart from that defeater) to confirm *p* over skeptical hypotheses concerning *p*" (2015, 60).

PLANTINGIAN DEFEATERS

In the previous chapter we briefly (and no doubt incompletely) surveyed the current literature on epistemic rationality and defeaters. In this chapter we will delve a little deeper into one of the accounts of defeaters that was briefly introduced in the previous chapter, that of Alvin Plantinga. The goal of this chapter is threefold: to present Alvin Plantinga's views about defeaters; to examine Jonathan Kvanvig's claim that, in *Warranted Christian Belief*, Plantinga turned his theory of knowledge into a defeasibility theory; and to present three recent criticisms of Plantinga's views about epistemic defeat, those of Jonathan Kvanvig, of Daniel Johnson, and of Max Baker-Hytch and Matthew Benton. I begin by laying out Plantinga's proper functionalism and his Evolutionary Argument Against Naturalism, since a previous general understanding of them seems to be required for an adequate comprehension of the remainder of the paper. With these preliminary matters out of the way, I present Plantinga's views about defeaters, briefly examine Kvanvig's claim that Plantinga has turned his theory of knowledge into a defeasibility theory, and conclude with a brief presentation of the aforementioned criticisms of Plantinga's views about epistemic defeat.

By looking more closely at one of the most ambitious attempts to provide a unified and complete account of epistemic defeasibility, I believe we will acquire a better understanding of how defeaters work and will thus be better positioned to evaluate the claims found in the literature to the effect that cognitive science of religion shows us that theism is irrational. As a consequence, our goal here is not to endorse or defend Plantinga's views of defeaters, but to examine his account with the goal in mind of deepening our understanding of an important contribution to our understanding of defeasibility in order to perhaps gain greater insight into the nature of defeaters and the questions and problems that might remain to be solved, in a way that can perhaps be helpful as we explore the question of rationality of theistic beliefs in light of current cognitive science.

PRELIMINARIES

Talk of defeaters entered the epistemological scene in Gettier's wake. One of the most important earlier attempts at solving the Gettier problem involved adding a fourth clause to

the three clauses of the traditional definition of knowledge (*i.e.*, *p* is true; *S* believes that *p*; *S*'s belief that *p* is justified) that would take into account the intuition that "justification that is good enough for knowledge withstands the addition of true beliefs to the agent's doxastic system" (de Almeida and Fett, 2015, 4) Although that fourth clause came in a variety of different formulations, it essentially said that *S*'s justification for believing *p* is undefeated.[73]

Although initially drawn from legal terminology,[74] defeaters, therefore, entered the philosophical scene with the promise of resolving one of the main contemporary epistemological conundrums. The use of defeaters in epistemological talk, however, was soon appropriated by other accounts of warrant other than the one offered by proponents of the defeasibility theory.[75] As Michael Bergmann noted, all the main proponents of epistemological externalism put forward one internal condition, namely a no-defeater one,[76] as necessary for warrant. Alvin Goldman adds as a necessary condition for warrant of *S*'s belief that *p* that *S* not believe that *p* is undermined (Bergmann, 1997, 405). Robert Nozick includes as a condition for warrant of *S*'s belief that *p* that *S* not believe that *p* does not track the truth (Ibid.). And Alvin Plantinga adds to his proper functionalist account of warrant that *S*'s defeater system is functioning properly. Plantinga and Bergmann himself both developed detailed accounts of defeaters – how they work, what types of defeaters there are, etc. And,

[73] Or as Feldman put it, "there is no true proposition t such that, if S were justified in believing t, then S would not be justified in believing p." Feldman, Richard (2003) *Epistemology*. Upper Saddle River: Prentice Hall, p. 34. Michael Sudduth summarizes the generic idea as "that a person S knows p only if there is no true proposition, d, such that if S were to believe d (or d were added to S's evidence for p), S would no longer be justified in believing p. In other words, the existence of certain unpossessed evidence prevents a person from actually knowing p if this unpossessed evidence would result in a loss of justification were the person to acquire the evidence, be aware of it, or recognize it." Sudduth, Michael (2008) Defeaters in Epistemology.

[74] See Sudduth, Michael (2008).

[75] Jonathan Kvanvig provides a good summary of the centrality of the notion of epistemic defeat for contemporary epistemology: "The concept of epistemic defeat, or some surrogate for it, is essential for any fallibilistic epistemology. If knowledge requires infallibility, then the epis-temic grounds of belief have to be strong enough that no further information could be made available to the cognizer to undermine these grounds of belief. When knowledge requires no such infallibility, however, grounds of belief can be undermined by further information, information that defeats the power of the original information to put one in a position to know that the claim in question is true. Even if some combination of conditions for knowledge are sufficient for truth, if there is a nonpsychological condition for knowledge that is not sufficient for truth, that condition will need to appeal to some concept of defeat (or a surrogate of it)" (Kvanvig, 2007, 104).

[76] According to Bergmann, this non-defeater condition (NDC) "is satisfied by S's belief that p if and only if S does not believe (and would not upon reflection) that her belief that p is defeated." In (1997, 407).

more recently, Jennifer Lackey and Sanford Goldberg have developed accounts of normative defeaters (defeaters constituted by doubts a subject ought to have given the evidence she has).[77]

The purpose of this chapter is to present one particular account of defeaters, the one developed by Alvin Plantinga, compare his account with that developed by defenders of the traditional defeasibility theory, as well as examine three recent criticisms of Plantinga's account, those by Jonathan Kvanvig, Daniel Johnson, and Max Baker-Hytch and Matthew Benton.[78] Plantinga has developed an account of defeaters that includes three principles of defeat and a typology of defeaters. His writings on the subject are, however, dispersed in several articles and books. My aim is to bring together the main things he has written on the subject and then examine the suggestion that he has turned his theory of knowledge into a defeasibility theory and present some recent criticisms of some aspects of his account of defeat. But since this defeasibility system is in many aspects connected to two important epistemological ideas defended by him (his proper functionalism and his Evolutionary Argument against Naturalism, henceforth the EAAN),[79] a brief exposition of them seems necessary before we can proceed to fulfill our main objectives.

Proper Functionalism

[77] As Lackey defines it, "A normative defeater is a doubt or belief that S ought to have and that indicates that S's belief that p is either false or unreliably formed or sustained. Defeaters in this sense function by virtue of being doubts or beliefs that S should have (whether or not S does have them) given the presence of certain available evidence. For example, suppose that Bill believes that the President is currently in Chicago, but then reads in The New York Times that the President is currently in China. If Bill continues to hold his original belief with no reason for doubting the report in the newspaper, it may be argued that even if the President is in fact in Chicago, Bill does not know this because there is evidence available to him that defeats his knowledge (justification/warrant). The underlying thought here is that certain kinds of doubts and beliefs—either that one has or should have—contribute epistemically unacceptable irrationality to doxastic systems and, accordingly, knowledge (justification/warrant) can be defeated or undermined by their presence." Lackey, Jennifer (2008) *Learning from Words*, p. 45. Goldberg offers the following definition of normative defeaters: "There are circumstances under which one or more propositions which S should believe (but doesn't) defeat the justification for S's belief that p." In Goldberg, Sanford (2015b) *The Asymmetry Thesis and the Doctrine of Normative Defeat*, https://www.youtube.com/watch?v=P2ZJNZIVlmY (9:20min.).

[78] The two last criticisms are directed to the notion of defeaters in general, not just to Plantinga's account.

[79] Although Plantinga says that the EAAN "does not presuppose this [proper functioning] or any other specific epistemology" (2002b, 205n2) his principles of defeat are formulated in the context of responding to objections to the EAAN. Thus, a general introduction to the EAAN seems to be very useful if not required for adequate comprehension of what will follow.

Alvin Plantinga has proposed an account of warrant[80] in which the proper function of our cognitive faculties plays a central role. According to proper functionalism, in order to have warrant, a belief has to be produced by cognitive faculties that are functioning properly (that is, free from malfunction or dysfunction). Proper function involves the notion of design plan (by evolution or God, or both): the cognitive faculties work properly if they are working as they should work. Like other of our human organs, there are ways in which our cognitive faculties are supposed to work. A properly functioning human heart, for instance, should beat about 50-80 times a minute when we are at rest. Moreover, the design must be a good one in the sense that the purposes of the design plan will be achieved. In the case of the human heart, its purpose is to pump the blood. Likewise, cognitive faculties can function properly or malfunction. The main component of a proper functionalist account of warrant is, therefore, that the cognitive faculties are functioning properly, subject to no significant dysfunction.

But there is more. Our human organs are designed to work in certain environments. We – to use some of Plantinga's examples – can't breathe under water; our muscles atrophy in zero gravity; we can't get enough oxygen at the top of Mt. Everest (Plantinga, 2015, position 686 of Kindle). Likewise, cognitive faculties will function properly only in environments for which they were designed. They wouldn't work well, for instance, to use another of Plantinga's examples, in an environment in which "a certain radiation impedes the function of memory" (Ibid.).

In addition to cognitive faculties working according to a good design plan and in an environment for which they were designed to work, the purpose of cognitive faculties functioning properly is to produce true beliefs. Such cognitive faculties must, therefore, be aimed at producing true beliefs. And since the design plan is a good one, there has to be a high probability that beliefs produced according to the plan will be true.

Thus, as Plantinga summarizes his account of warrant,

> *A belief has warrant for a person S only if that belief is produced in S by cognitive faculties functioning properly (subject to no*

[80] Warrant is that, whatever precisely it is, which makes the difference between knowledge and true belief.

> *dysfunction) in a cognitive environment that is appropriate for S's kind of cognitive faculties, according to a design plan that is successfully aimed at truth* (Ibid.).

The EAAN

According to orthodox Darwinian evolutionary theory, we human beings are the product of a biological evolutionary process driven mainly by two mechanisms: natural selection and random genetic mutation. As a result of the first mechanism, the genetic mutations that are maladaptive and don't enhance fitness are discarded and those that have adaptive value and enhance fitness are preserved as part of the genome. Our cognitive faculties are, therefore, the product of this process. But if naturalism, the thesis – as Plantinga defines it – according to which there is no God or anything like God, is true, this evolutionary process was unguided, i.e., there was no supernatural guidance of this process with the aim of producing the sort of beings that we are. But if this is so why should we think that this process would have produced beings with reliable cognitive faculties? After all, natural selection will favor cognitive faculties and processes that result in adaptive behavior, but whether true beliefs will be formed is irrelevant for this process. Our behavior could be adaptive but our beliefs false as frequently as true. Thus, an unguided evolutionary process would not give us reason to believe that our belief-producing processes are for the most part reliable. The naturalist, then, by believing that the evolutionary process was unguided, is not in a position to affirm that it is probable that this process would result in creatures with for the most part reliable cognitive faculties that would produce a preponderance of true beliefs.

With this picture of the evolutionary origins of our cognitive faculties in place, it is easy to see, argues Plantinga, that the probability of our cognitive faculties being reliable, given naturalism and evolution, is low. And if I believe in naturalism and evolution and come to see that this probability is low, I have a defeater for my belief that my cognitive faculties are reliable (a Humean defeater in Plantinga's terminology, as we shall see).

In *Where the Conflict Really Lies* (2012), Plantinga presents schematically his latest version of the argument as follows [R being reliability of our cognitive faculties; N being naturalism; and E being evolution]:

(1) P(R/N&E) is low;

(2) anyone who accepts (believes) N&E and sees that P(R/N&E) is low has a defeater for R;

(3) anyone who has a defeater for R has a defeater for any other belief she thinks she has, including N&E itself;

(4) if one who accepts N&E thereby acquires a defeater for N&E, N&E is self-defeating and can't rationally be accepted – conclusion: N&E can't rationally be accepted (Ibid., 345).

PLANTINGIAN DEFEATERS

From what we saw above it seems clear that the notion of defeat has been crucial to Plantinga's work in epistemology. Plantinga has availed himself of expressions such as: warrant defeaters, proper-function-rationality defeaters, Humean defeaters, purely alethic rationality defeaters, potential defeaters, defeater-deflectors, defeater-defeaters, neutralizing defeater-defeaters, intrinsic defeater-defeaters, extrinsic defeater-defeaters, partial defeaters, and "optimistic overriders," among others. No doubt an impressive panoply of defeaters! In what follows we will see what he means by each of these expressions, as well as see three principles of defeat he has developed, among other important issues related to defeaters.

Plantinga accepts John Pollock's distinction between rebutting and undercutting defeaters. The former are reasons you have for taking a belief *b* you hold to be false and, thus, you have to give up that belief. Here is one example: You see what looks like a sheep in the field and form the belief that there is a sheep in the field. But the owner of the field, a trustworthy man, comes and tells you that there is no sheep in the field, but that there is a sheepdog that from that distance is likely to look like a sheep. If you are rational, you will no longer hold the belief that there is a sheep in the field. The owner's testimony about the sheepdog constitutes a defeater for your belief that there is a sheep in the field.[81]

[81] Example found in *Where the Conflict Really Lies*, p. 164.

Undercutting defeaters, on the other hand, are not reasons to believe a belief *b* you hold is false, but, instead, that you don't have reason for holding it. One example: You see someone leaving the house across the street and form the belief that Paul is leaving the house. But you are told that Paul's twin brother, Peter, arrived recently at Paul's house. This information is not enough to give you a rebutting defeater for your belief that Paul was leaving the house. However, it compels you to withhold it, to be agnostic about whether or not Paul was the one leaving the house.[82] Plantinga defends that undercutting defeaters come in degrees. They may just reduce the firmness by which you hold the belief instead of compelling you to completely withhold it. For example, suppose Lucas and Rogel tell me they overheard Joaquim Clotet say, as he was leaving his office, that he has plans to step down as the President of PUCRS by the end of year and that Professor Claudio has been chosen as his successor and has agreed to it. I believe that is true. But later on Lucas tells me that he didn't actually overhear it but was just relying on Rogel's testimony. This new information partially undermines the reason I have for believing that Professor Claudio will succeed Clotet. I may continue to believe that this is the case, but less firmly.[83]

Another important distinction is the one between warrant and rationality defeaters. Warrant, as we saw above, is the quality that distinguishes knowledge from mere true belief. Warrant defeaters are "circumstances that result in my belief's failing to have warrant in a state of affairs where it otherwise would have it" (2012, 166). In the barn case, for instance, your belief that there is a fine barn in front of you is not knowledge because most of the barns in the area are fake ones and that you form the true belief that that one is a fine barn results from sheer good luck. There is no failure of rationality in this case. The defeater you have here is for the warrant you have for your belief that that's a fine barn. The defeater here need not be a belief, but just some feature of the environment. The same is not true of rationality defeaters, as in the sheepdog and Paul-Peter cases above, where the defeaters you acquire come by way of beliefs (alternatively, it could come by way of an experience). Of

[82] Example found in *Where the Conflict Really Lies*, p. 165.

[83] Adaptation from an example provided by Plantinga in *Where the Conflict Really Lies*, p. 252.

course, the subject need not become aware of the defeater for it to defeat her warrant for a certain belief. And all rationality defeaters are also warrant defeaters.[84]

Defeaters are relative to our noetic structure, the rest of what we know and believe. In the sheepdog case, the testimony of the owner of the field gave me a defeater for my belief that there was a sheep in the field because I already believed he was an honest man. If, instead, I believed he was a contrarian epistemologist who loved to provide false defeaters for beliefs of other people just for the fun of it, I wouldn't have acquired a defeater. As a consequence, what is a defeater for me in a given circumstance may not be a defeater for you.

Sometimes a false belief can be a defeater. Plantinga (2000, 368) illustrates this by supposing that a certain theist believes that if theism is true, then there couldn't be any coherent Freudian projective theories – which claim that theistic belief is the product of wish-fulfillment. And the theist ends up believing that there indeed are coherent Freudian projective theories. This provides a defeater for her theism. But the belief that theism cannot be true if there are coherent Freudian projective theories is false. So her theistic belief was defeated by a false belief.

And can a defeater be acquired irrationally? Plantinga's answer is yes. There are, in Plantinga's view, two different forms of epistemic rationality: there is internal and there is external rationality. While the former is, as Plantinga put it, "a matter of proper function 'downstream from experience,'" of appropriate response to experience, the latter is "a matter of the proper function of the sources of experience" (2012, 365). External irrationality can arise by different kinds of cognitive malfunction. Suppose, to adapt another of Plantinga's examples, that I, due to the development of a paranoid condition, start believing that a good friend of mine, someone whom I have trusted and respected greatly

[84] This distinction between warrant and rationality defeaters is formulated in *Warranted Christian Belief* (2000) and *Where the Conflict Really Lies* (2012). In *Naturalism Defeated* (1994) Plantinga has the following to say about this distinction and what differentiates his views of defeaters from those of Peter Klein:

> *Contrast Peter Klein, 000000. Klein's conception of a defeater, so far as I can see, is of something that defeats the warrant a belief enjoys. This is a perfectly sensible way to think about defeaters: more exactly, it is perfectly sensible to think that there are defeaters of this sort. My concern, however, is with defeaters that defeat the rationality of a belief, not its warrant. Both kinds of defeaters are important; rationality defeaters are what are relevant in this context, in which we are thinking about the rationality of a certain belief (N&E), not about its warrant.*

for a long time, wants to destroy my career by spreading lies about me. Can belief D – she is trying to destroy my reputation – be a defeater of my belief B – that she is a good and trustworthy friend? Belief D is acquired by cognitive malfunction and is, therefore, externally irrational. And, at the same time, internal rationality requires me to give up the belief B. Hence, D, despite being acquired irrationally, functions as a defeater for my belief B.

Sometimes I can believe a defeater but I don't see its bearing on the defetee, and unless and until I see the connection between the defeater and the defeatee, the former does not produce its defeating effect. It lies dormant as a *potential defeater* until this occurs. Here is one of the cases Plantinga uses to illustrate how potential defeaters work:

> *Frege once believed that*
>
> *(F) For every condition or property P, there exists the set of just those things that have P.*
>
> *Bertrand Russell wrote him a letter, pointing out that (F) has very serious problems. If it is true, then there exists the set of non-selfmembered sets (because there is the property or condition of being non-self-membered). This set, however, inconsiderately fails to exist. That is because if it did exist, it would exemplify itself if and only if it did not exemplify itself; that is, it would both exemplify itself and fail to exemplify itself, which is wholly unacceptable behavior for a set. Before he realized this problem with (F), Frege did not have a defeater for it. Once he understood Russell's letter, however, he did; and the defeater was just the fact that (F), together with the truth that there is such a condition as being non-self-membered, entails a contradiction* (2000, 361).

In *Warranted Christian Belief* (2000) Plantinga stipulates the following definition of defeater (*simpliciter*):

> *(D) D is a defeater of B for S at t if and only if (1) S's noetic structure N (i.e., S's beliefs and experiences and salient relations among them) at t includes B, and S*

> *comes to believe D at t, and (2) any person (a) whose cognitive faculties are functioning properly in the relevant respects, (b) whose noetic structure is N and includes B, and (c) who comes at t to believe D but nothing else independent of or stronger than D would withhold B (or believe it less strongly)* (2000, 362).[85]

But there is a problem here. According to this definition, a "belief D is a defeater of B for you if proper function requires giving up belief B when you acquire D" (2000, 362). But then proper function could require that one believes certain things not because they are true but because they contribute to survival or produce psychological comfort. Suppose Bill is seriously ill, but nonetheless believes, on the basis of wish-fulfillment, that he will soon recover. Suppose further he is apprised of some statistics to the effect that people in his condition are unlikely to recover. Perhaps proper function requires in this case that Bill continues to believe that he will recover (2000, 363). Thus, we need, according to Plantinga, the notion of a purely epistemic defeater:

> *(D*) D is a purely epistemic defeater of B for S at t if and only if (1) S's noetic structure N at t includes B and S comes to believe D at t, and (2) any person S* (a) whose cognitive faculties are functioning properly in the relevant respects, (b) who is such that the bit of the design plan governing the sustaining of B in her noetic structure is successfully aimed at truth (i.e., at the maximization of true belief and minimization of false belief) and nothing more, (c) whose noetic structure is N and includes B, and (d) who comes to believe D but nothing else independent of or stronger than D, would withhold B (or believe it less strongly)* (2000, 363).

The purely epistemic defeater, then, would not defeat beliefs formed by processes not aimed at truth (but at survival or psychological comfort). The point, says Plantinga, is that "D could be a purely epistemic defeater of B even if proper function requires the maintenance

---

[85] *Warranted Christian Belief*, p. 362. A similar definition is presented in *Naturalism Defeated* (1994, p. 33), and Plantinga adds that

> *We could also put it in terms of the human design plan: given noetic structure N and new belief D, that design plan requires the deletion of B from S's noetic structure. We could also put it like this: D is a defeater of B for you if your noetic structure includes B at t; at t you come to believe D; and rationality requires that if you continue to believe D, you will cease believing B.*

of B, in S's noetic structure, despite the formation of D; that can occur if the processes maintaining B are not aimed at truth" (2000, 363). With this distinction between defeaters *simpliciter* and purely epistemic defeaters in place, a Freudian who argues that theistic belief is the product of wish-fulfillment could defend that, even though the evil in the world doesn't give the theist a defeater *simpliciter* for her theistic beliefs, it does provide her with a purely epistemic defeater.[86]

We mentioned in the section on the EAAN that the naturalist who recognizes that P(R/N&E) is low acquires a Humean defeater. The defeater here is a rationality defeater. To illustrate how such defeaters work in the case of the naturalist who sees that P(R/N&E) is low, suppose

> *there is a drug – call it XX – that destroys cognitive reliability. I know that 95 percent of those who ingest XX become cognitive unreliable within two hours of ingesting it; they then believe more false propositions than true. Suppose further that I come to believe both that I've ingested XX a couple of hours ago and that P(R/I've ingested XX a couple of hours ago) is low; taken together, these two beliefs give me a defeater for my initial belief that my cognitive faculties are reliable* (2012, 342).

But in what sense does the person who took XX and became aware of its effects acquired a defeater for the belief that her cognitive faculties are reliable? Won't she be unable to stop carrying on with her daily affairs and assume her cognitive faculties are reliable? In *Reply to Beilby's Cohorts* (2002b), Plantinga calls the defeater *simplicter* (D) a proper-function rationality defeater and the purely epistemic defeater (D*) a purely alethic rationality defeater. In the case of the person who ingested XX, proper function, as in the case of wish-fulfillment, may dictate that she continues to assume that her cognitive faculties are reliable. She, in other words, does not obtain a proper-function rationality defeater, but if her truth-aimed processes were at work she would have a purely alethic rationality defeater. Her

[86] The theist, of course, would argue – as Plantinga argues later in *Warranted Christian Belief* – that the problem of evil does not constitute a defeater (*simpliciter* or purely epistemic) for theistic (and Christian) belief.

situation here is analogous to that of Hume, who concludes that most things – and the most important things – he believes are highly doubtful. Reason leads him to conclude that the things he normally – outside his study – believe to be true may not be true after all. That's not a comfortable situation (to put it mildly), as he himself confesses:

> *I am confounded with all these questions, and begin to fancy myself in the most deplorable condition imaginable, invironed with the deepest darkness, and utterly deprived of the use of every member and faculty.*[87]

But the same nature that prompts him to think about these things also leads him, after he leaves his study for a new round of a game of backgammon, to be relieved of such anguish:

> *Since reason is incapable of dispelling these clouds, nature herself suffices to that purpose, and cures me of this philosophical melancholy and delirium, either by relaxing this bent of mind, or by some avocation, and lively impression of my senses, which obliterate all these chimeras. I dine, I play a game of backgammon, I converse, and am merry with my friends.*[88]

The predicament of the reflective naturalist is the same of the person who took XX. She acquires a purely alethic rationality defeater once she realizes P(R/N&E) is low. But she – and the naturalist – will also acquire a proper-function rationality defeater for R – a Humean defeater – whenever she thinks about her cognitive situation (2002b, 211).

There are also partial defeaters, defeater-defeaters, and defeater-deflectors. The first are "defeaters that don't require withholding B but do require holding it less firmly" (2000, 362,

---

[87] The full quotation from Hume found in (2002b, 210) says:

> *Where am I, or what? From what causes do I derive my existence, and to what condition shall I return? Whose favour shall I court, and whose anger must I dread? What beings surround me? and on whom have, I any influence, or who have any influence on me? I am confounded with all these questions, and begin to fancy myself in the most deplorable condition imaginable, invironed with the deepest darkness, and utterly deprived of the use of every member and faculty.*

[88] Cited in 2002b, 210.

note 3). A defeater that defeats another defeater is a defeater-defeater. Some beliefs have a higher degree of warrant than potential defeaters for them, being thus intrinsically defeaters of such defeaters, reason why Plantinga calls such beliefs intrinsic defeater-defeaters. Plantinga illustrates what he means by the concept of intrinsic defeater-defeater with the following example:

> *I apply for a National Endowment for the Humanities fellowship; realizing I am not really qualified, I offer you five hundred dollars to write a glowing if inaccurate letter of recommendation. Perhaps, as they say, everyone has a price; as it turns out, yours is definitely more than five hundred dollars. You indignantly refuse, and write a blistering letter to the chair of my department. The letter mysteriously disappears from her office. One of the most respected members of the department, however, reports having seen me apparently trying to enter her office through a second-story window. I have means, motive, and opportunity. Further, I am known to have done this sort of thing before. But I clearly remember being on a solitary hike in the mountains the entire afternoon during which the letter disappeared. I believe that I did not remove that letter, and that belief has warrant for me* (2000, 371).

Here the belief provided by memory that he was hiking in the mountains when the crime took place is an intrinsic defeater of the defeater provided by the evidence that she committed the crime. There are also extrinsic defeater-defeaters: "a belief *r* that defeats a defeater *q* of a belief *p* distinct from *r*" (1986, 311). And neutralizing defeater-defeaters: as the name indicates, the defeater-defeater doesn't defeat the original defeater, but, instead, neutralizes its defeating potential (1994, 36-7). "What happens," writes Plantinga,

> *is that at t your noetic structure N includes B; then at t* you come to believe something D which is a defeater of B, so that you move to a noetic structure that includes D but not B; then at t** you*

> *learn (come to believe) something D* such that its addition to your noetic structure permits a move to a noetic structure that includes D, D* and B* (1994, 37).

He gives the following example of how neutralizing defeaters work:

> *I have heard somewhere that you can't swim and at t believe that; at t* I learn that you are a lifeguard, which (together with my belief that nearly all lifeguards can swim) gives me a defeater for my belief that you can't swim; but then at t** I learn that you are a Frisian lifeguard and that only half of the Frisian lifeguards can swim, which gives me a defeater for that defeater; but then at t*** I learn that you graduated from the famous lifeguarding school at Leeuwarden, all of whose graduates can swim, which gives me a defeater for that defeater-defeater; and so on; we can add to the series ad libitum. If the last member of the series is odd numbered, I will wind up rationally holding the original belief along with its defeater, the defeater for that defeater, and so on* (1994, 37).

Defeater-deflectors are "belief[s] I already hold such that as long as I hold [them] (and given my noetic structure) I can't rationally come to hold B" (2012, 260). Here is an example provided by Plantinga. In the sheepdog case we saw above, the testimony of the owner of the pasture to the effect that there is no sheep but there is a sheepdog in the field gives me a defeater for my belief that there is a sheep there. But suppose we add to the story that prior to my talking to the owner his wife tells me he has the habit of telling everybody that there is no sheep in the field even though there is one. Since I believe she is telling the truth, I don't get a defeater from the owner's testimony. Her testimony acts as a defeater-deflector.[89]

[89] Plantinga notes that this "characterization of a deflector is too broad, in that it assigns deflectorhood to almost any belief with respect to a belief that is sufficiently irrational, in the sense that a properly functioning human being with anything like a standard noetic structure will not form B. But then nearly any other belief I

The difference between defeater-defeaters and defeater-deflectors is that the former requires that one first have a defeater D for the belief B, while the latter prevents D from being a defeater in the first place (2011, 439, note 11). In the case of the EAAN, some critics of the argument have argued that it is possible for the naturalist to add something to his belief system so that there would be a belief deflecting the defeater provided by the realization by the naturalist that P(R/N&E) is low (2011, 435). Carl Ginet (1995), for instance, proposes that R itself is a defeater-deflector, *i.e.*, according to Ginet, the naturalist could say that one of the tenets of naturalism is that our cognitive faculties are reliable. But, according to Plantinga, "If this were sufficient for deflecting a defeater, there wouldn't be any probabilistic defeaters at all." The theist, for instance, could respond to the probabilistic argument from evil by claiming that G (God) is a defeater-deflector of the proposed defeater for her theistic beliefs – the claim that P(G/E) is low, where E stands for evil –, *i.e.*, by acknowledging that P(G/E) is indeed low, but adding that she also believes G, and, thus, that P(G/G&E) is high.[90]

These are the types of defeaters Plantinga has catalogued and employed in his epistemological work. But he has also developed some principles and conditions under which there can be defeat. These principles were developed in the context of Plantinga's responses to some objections to the EAAN. According to one of the objections, if N&E is a defeater for R because the P(R/N&E) is low, then it would also be the case that N&E is a defeater for any other proposition that would have a low probability given N&E, such as that *Juiz de Fora was at her glorious days known as the "Manchester of Minas"*, or that *the function of perspiration is to cool the body*. Thus this has been labeled *the perspiration objection*.[91] But since it is absurd to say that N&E is a defeater for those propositions, it is

---

hold is (on the above account) a deflector with respect to B. What has to be added is a clause roughly to the effect that a deflector belief D for a belief B must be such that in the relevant circumstances, if D were not present in my noetic structure, I would have formed B. True, this account involves a counterfactual, inviting the sort of grief to which analyses involving counterfactuals often succumb; but perhaps it is close enough for present purposes. In any event, I leave as homework the project of refining the account" (2012, 260).

[90] In *Content and Natural Selection*, Plantinga evaluates the potential of reductive materialism, nonreductive materialism, indicator semantics, functionalism, and teleosemantics as deflectors of the defeater constituted by P(R/N&E).

[91] This objection has been wielded against the EAAN by, among others, Talbott, William (2002).

likewise absurd to say that N&E is defeater for R. It seems the principle behind this argument is

> *(1) For any propositions A and B I believe, if B is improbable or inscrutable with respect to A (i.e., the right attitude towards the question of its probability with respect to A is agnosticism) then A is a defeater for B* (1994, 40).

But this principle is false. Plantinga shows that by way of an example. Suppose I believe

> *(2) You own an old Nissan,*
>
> *and I also believe*
>
> *(3) You own a Japanese car* (1994, 40).

Although (2) is improbable with respect to (3), (3) is not a defeater for (2). If it were, it would be irrational to continue to believe (2) after realizing that it is improbable on (3). But clearly there are circumstances in which a belief A is improbable with respect to B and B is in fact a defeater for A. The difference, according to Plantinga, is that in the first case the warrant (3) has for me is derivative from the warrant (2) has for me. (3) gets the warrant it has by virtue of being inferred from my knowledge of (2). With this in mind, he formulates his first principle of defeat:

> *(First Principle of Defeat (FPD)) If S rationally believes that the warrant a belief B has for him is derivative from the warrant a belief A has for him, then B is not a defeater, for him, of A* (1994, 41).

Another sort of objection that has been proposed against the EAAN is that austere theism[92] is a defeater for R. This argument has come in several varieties. According to one of them, in the same way the naturalist acquires a defeater for R when she sees that P(R/N&E), the theist acquires the same type of defeater once she sees that P(R/A&E), where A stands for austere theism, is low. But the problem here, says Plantinga, is that suppose I know that you

[92] Austere theism, unlike theism *simplicter*, does not include the proposition that we have been created in the image of God, which is what differentiates the theist's and the naturalist's epistemic condition with respect to the EAAN.

own an old Nissan and infer from it both that you own an old car and that you own a Japanese car. But suppose I realize that that you own a Japanese car is improbable given that you own an old car and conclude that the latter is a defeater for the former. Clearly, defeaters don't work that way. What has gone wrong here is that the beliefs that you own an old car and that you own a Japanese car obtain their warrant from the same belief (my knowledge that you own an old Nissan). The warrant of the inferred beliefs are derivative from the belief that you own an old Nissan. In the case of theism, she rationally believes the warrant she has for austere theism is derivative from the warrant she has for theism *simpliciter* and "under those conditions the former is a defeater for a belief A – R, for example – only if the latter is. But it isn't" (1994, 46). From this Plantinga derives the following principle:

> *(Second Principle of Defeat) (SPD) If S rationally believes that the warrant, for him, of a belief B is derivative from that of a belief A, then B won't be a defeater, for him, for any belief C unless he rationally believes that A is a defeater for C* (1994, 45).

A third category of objections to the EAAN is the one that became known by the slogan *why can't the naturalist just add a little something*? And it has been defended by Ginet (1995) and O'Connor (1994), among others. According to Ginet, as we saw above, if the theist can escape irrationality by believing something more than mere austere theist, why can't the naturalist just add a little something too (such as that our cognitive faculties are reliable or that the naturalist has won the evolutionary lottery)? Plantinga, however, doesn't think that the naturalist that adds that her cognitive faculties are reliable to N, call it N^+, is doing the same thing than the theist that goes beyond austere theism. For "the warrant austere theism has for the theist is derivative from the warrant theism has for her; but it is not the case that the warrant N&E has for the naturalist is derivative from the warrant N^+ has for him" (1994, 50). Ginet's proposal can be constructed as a claim that N^+ is a neutralizing defeater for the defeater of R provided by N&E. But the problem here, says Plantinga, is that the warrant the conjunction N&E has for us is derivative from the warrant of its conjuncts. But since one of those conjuncts is the defeatee itself, a successful defeater-defeater doesn't materialize. Thus a third principle:

(Third Principle of Defeat) (TPD) If D is a defeater of B for S, then for any belief B of S, if S rationally believes that the warrant B* has for her is derivative (wholly or partly) from the warrant B has for her, then B* is not a defeater-defeater, for S, of D* (1994, 51).

PLANTINGA'S DEFEAT SYSTEM AND THE TRADITIONAL DEFEASIBILITY THEORY

According to Jonathan Kvanvig, in *Warranted Christian Belief* (2000), Plantinga "turns his theory into a defeasibility theory" (2005). What does he mean by that? Does he mean that Plantinga has joined the ranks of the early Lehrer, Klein, Pollock, Swain, and De Almeida & Fett, among others (but, to be sure, probably not many others…)? In principle, nothing should prevent him from doing so. As Marshall Swain[93] – and Plantinga himself[94] – note, the defeasibility theory can be accommodated into virtually any theory of warrant. One may suspect, however, that that is not what Kvanvig meant in his commentary. One may think that he probably only meant that Plantinga has added a defeater system to his theory of knowledge, that he added a no-defeater condition to it. Still, there is some evidence in favor of the view that Kvanvig could have meant that Plantinga has indeed turned his theory of knowledge into a defeasibility one. So let us pursue this question a little further.

The first problem with the idea that what Kvanvig meant was that Plantinga added a no-defeater condition is that he had done so long before *Warranted Christian Belief*. In *Respondeo* (1996), Plantinga's response to the commentators on his theory of warrant in the book edited by Kvanvig (1996), devotes a great deal of space criticizing the defeasibility theory as developed by Klein. Plantinga's statements regarding the addition of a defeater system to his theory of warrant could be found already in *Warrant and Proper Function*

[93] As Swain (1996, 134) puts it,

> *It is important to note that the indefeasibility strategy is not essentially linked with any particular account of epistemic justification. One can be an internalist, or an externalist with respect to the nature of justification and still hold that knowledge is indefeasible justified true belief. Similarly, one can be a foundationalist or a coherentist. Virtually all of the theories of justification considered and rejected by Plantinga as neither necessary nor sufficient for warrant could, with appropriate modifications, provide the analysis of justification required for the defeasibility account of knowledge.*

[94] Plantinga cites and endorses Swain's remarks about the defeasibility theory being compatible with virtually any theory of warrant. In Plantinga, Alvin, 1996, 318.

(1993, 40-42), which led Michael Bergmann to add Plantinga to the ranks of proponents of a no-defeater condition. In fact, Bergmann reports that Plantinga confirmed to him, in conversation, that "a necessary condition of warrant for S's belief that p is that S *not believe* that her belief that p is defeated" (1997, 405-6; my emphasis). This certainly is not enough to make Plantinga a defeasibilist, as he is defending the absence of internal defeaters or overriders, not propositional defeaters, which is what characterizes the view defended by proponents of the defeasibility theory.

One important piece of evidence in favor of the thesis that Plantinga has indeed moved his theory of knowledge in the direction of that defended by proponents of the defeasibility theory is Michael Sudduth's formulation of what he calls "the no-defeater condition":

> [ND] Given any person S, S's belief B (held to some degree n) is warranted only if S does not have an undefeated defeater for B) (1999, 171).

Here Plantinga's "no-defeater condition" is formulated in terms of propositional defeaters. And according to Sudduth,

> *Plantinga has informed me in correspondence that a person can have an undefeated defeater without believing that she has it, and that would be enough to defeat warrant. Lastly, in more recent correspondence Plantinga has suggested that one can believe that one has a defeater for a belief B without actually having one, in which case B could still be warranted, though perhaps not to the same degree* (1999, 184, note 12).

Is this statement consistent with Plantinga's stated position vis-à-vis the defeasibility theory? As we have seen, Plantinga makes a distinction between rationality and warrant defeaters. And already in *Naturalism Defeated* (1994), as we saw in note 18, in explaining what differentiates his views about defeaters from those of Peter Klein, he said that both warrant and rationality defeaters are important. Still, the matter is not as clear as one might wish.

Warrant defeaters are important, then. They can defeat one's warrant. But a defeasibility condition is not included in Plantinga's definition of warrant,[95] as we saw in the first section.

FOUR OBJECTIONS

Baker-Hytch/Benton's objection

Max Baker-Hitch and Mathew Benton's criticisms of Alvin Plantinga's account of defeat emerge from their more general criticism of the traditional notion of epistemic defeasibility. They formulate their objections to what they call "defeatism" as an improvement upon the criticism leveled by Maria Lasonen-Aarnio, in her article *Unreasonable Knowledge*. We will explore Lasonen-Aarnio's criticism in detail in the next chapter. It suffices to say now that she defends, on the basis on some cases of rationality defeat that she regards as not involving knowledge defeat, the thesis that rationality – or, as she prefers, reasonableness –, though connected to knowledge in important ways, it is not necessary for knowledge. As a result, while subjects who retain beliefs in cases of rationality defeat no longer believe rationally, they do not necessarily lose knowledge.

Baker-Hitch and Benton propose to go beyond Lasonen-Aarnio's contribution "by sketching an account based on the knowledge norm of what makes belief 'unreasonable' in some of cases of putative defeat" (2015, 40). They argue that both internalists and externalists have difficulty accommodating the notion of defeat into their favored epistemology. The difficulty for internalists comes from their conception of defeat in terms of probability-lowering. Externalists, on the other hand, have usually included an internalist no-defeater condition on knowledge. Baker-Hitch and Benton see this, however, as producing a "gerrymandered picture of knowledge (or justification)" (2015, 40)[96] with no connection with what

[95] "*A belief has warrant for a person S only if that belief is produced in S by cognitive faculties functioning properly (subject to no dysfunction) in a cognitive environment that is appropriate for S's kind of cognitive faculties, according to a design plan that is successfully aimed at truth."* Plantinga (2015, posição 686 do Kindle).

[96] As they put it, "For externalists about justification, however, simply adducing an internalist no-defeat condition on knowledge (or justification) is not an attractive option. Doing so results in a gerrymandered picture of knowledge (or justification) that includes an internalist defeat condition bearing no connection with the deep structural feature that they, qua externalists, take to be characteristic of knowledge (*e.g.* safety, sensitivity, track-record reliability, aptness)." [Ibid.]

characterizes the externalist view about knowledge. They explore ways in which purely externalist accounts of defeat could be constructed and find all of them unsatisfactory.

Among the externalist views that they explore is the proper functionalist accounts of Alvin Plantinga and Michael Bergmann. They assess the idea that defeat cases should be seen as cases in which properly functioning noetic systems respond to defeating evidence with suspension of belief. They claim that the proper functionalist account doesn't seem to have the resources to deal with the notion of normative defeat, *i.e.*, it doesn't seem to be able to explain why suspension of belief is the proper response in face of a merely potential defeater. The proper functionalist has to provide an account of why giving up a belief in face of a defeater is the proper way for a noetic system to function. This, Baker-Hitch and Benton suggest, could be explained in counterfactual terms – what that agent would counterfactually do were she apprised of the defeater – or statistically – proper response to defeaters is formulated in terms of what most cognitive agents in fact do. They argue that whatever option the proper functionalist chooses, she faces serious difficulties in explaining defeat. When explained in counterfactual terms, the proper functionalist view of defeat "loses normative universalizability" (2015, 54) *i.e.*, it becomes restricted to the manner each individual responds in her own particular way to defeating evidence. While it may be true that if I were apprised of defeater *d* for the evidential support e that I have for *p*, I would drop belief that *p*, it doesn't mean that another subject will have the same counterfactual disposition to drop *p* when in the same evidential situation that I am. And the appeal to the statistical norm, on the other hand, faces the problem that it "makes one's view hostage to empirical fortune:"[97] if, for instance, most agents do not suspend judgment in face of testimony to the effect that a wall that seems white to me is being irradiated by red lighting, then the proper response is to continue believing that the wall is red.

Kvanvig's objections

Jonathan Kvanvig has made a distinction between two approaches to epistemic defeat: the front door and the back door approaches. The first theory holds that our understanding of

[97] 2015, 55.

how epistemic defeat works must begin with a picture of propositional relations dependent upon evidential relations. Such relations hold prior to the defeater being added to the subject's noetic system, at its front door, so to speak. By contrast, the second theory accounts for defeat in terms of appropriate epistemic responses to the addition of new information to a noetic system. Alvin Plantinga's proper functionalism about defeasibility is taken by Kvanvig to be the best-developed example of such a theory. Once a belief enters one's noetic system, it is the proper functioning of such system that will determine whether it will play the role of a defeater and which belief will play the role of defetee, i.e., which belief will be ejected from the system. Kvanvig calls it a back door approach since, under this account, "we insert the defeater into the noetic house, and see which belief gets expelled out the back door" (2007, 111). Thus, while the front door approach characterizes defeat by the antecedent propositional relations of the defeater before it enters the noetic system, the back door characterizes it by the sort of response that would be prescribed by a noetic system as information is added to it.

As Kvanvig sees it, however, back door approaches are unable to accommodate the notion of defeater-defeaters. The idea here is that, since in a back door approach defeat is regulated by what the properly functioning noetic system prescribes, the same must be true of defeater-defeaters. For a belief to count as a defeater-defeater it must be present in the noetic system as the defeater enters it. But if so, no belief will be expelled and, hence, the defeater is not in fact a defeater. And since, as we saw above, defeater-defeaters are an important component of Plantinga's approach to defeasibility, if Plantinga's epistemology cannot make room for this kind of defeater, this seems to be a problem for his attempt to formulate a comprehensive theory of defeasibility.[98]

[98] And, in fact, as Kvanvig notes, no adequate account of defeat can forgo defeater-defeaters, "since the possibility of such follow straightforwardly from an appropriate understanding of the fallible character of reasons for belief: defeaters are no more infallible reasons to abandon belief than is evidence an infallible reason to hold a belief." Ibid., p. 111.

Kvanvig defends that another problem with Plantinga's approach to defeat – and, in fact, with all doxasticist theories of epistemic justification, as Kvanvig calls them[99] – emerges when we consider the Quine/Duhem problem:

> [Quine and Duhem's] point is that when we test a hypothesis and get results in conflict with the hypothesis, the existence of auxiliary hypotheses involved in the testing prevents the test from forcing the conclusion that the tested hypothesis is false. Instead, there is a variety of rational responses to an anomalous experimental result. As a consequence, one may expect properly functioning noetic structures to display no single response to the introduction of a defeater. Instead, there can be a variety of changes displayed by systems that are both reliable and properly functioning.

Thus, if Kvanvig is right, proper functionalists and other doxasticists face difficulties with defeater-defeaters and with the Quine/Duhem problem. Propositionalists, on the other hand, would be naturally able to accommodate defeater-defeaters. For, as Kvanvig put it, "Where *dd* is a defeater defeater of *d*, the conjunction of any evidence e conjoined to *d* does not justify *p*, but the conjunction of *e* plus *d* plus *dd* yields at least as much justification for *p* as provided by *e* itself" (2007, 119). And, unlike doxasticism, propositionalism would be able to offer a satisfactory account of defeat even in face of the Quine/Duhem problem. In the circumstances affected by the Quine/Duhem problem, the result that would be delivered by a back door approach would be paralysis. In a propositional or front door approach, on the other hand,

> [E]ven if d is a defeater of the p|e relation (where e is the evidence for p), it need not be a defeater of the p&r|e relation. That allows rational adjustments to a system of beliefs in response to learning d that don't require abandoning p. All learning d requires (on the assumption that there are no restorers present) is

[99] Doxasticist theories include proper functionalism, reliabilism, and virtue theory.

> that some evidentially suitable adjustment is made, one of which is abandoning p, but not the only one (2007, 119).

Johnson's objection

The general lines of Daniel Johnson's objection have common elements with both Kvanvig's and Benton/Baker-Hitch's objections. According to Johnson, proper functionalism has serious difficulties accommodating certain defeating experiences (Kvanvig's objection). And the defeaters in question here are normative defeaters (Benton/Baker-Hitch's objection).[100]

The core idea is presented by way of the following case:

> I am on a hike in the mountains, and read in my guidebook (which I justifiably believe is reliable) that you cannot see any lakes from the peak. I thereby form the justified belief that I won't be able to see any lakes from the peak. I then, an hour later, reach the peak, look down, and see what is obviously a lake. For whatever reason, though, I fail to put two and two together and do not come to believe that there is a lake there. I continue to believe that you cannot see any lakes from the peak and do not regard my belief that you cannot see any lakes from the peak to be defeated. (There is nothing really obscure or artificial about this case; it is the sort of thing we experience often. The lake experience just doesn't register on me, though I form many other beliefs normally while at the top of the peak. When I come down from the peak, perhaps a stranger asks me whether you can see any lakes from the peak. I blithely answer "no" because I'm still holding rather ridiculously to the testimony of the guidebook. Perhaps I have a friend who accompanied me to the peak, and when I tell the stranger that you can't see any lakes from the peak, my friend stares at me incredulously. He asks with a laugh, "Dan, are you

[100] Johnson's objection is directed specifically against Michael Bergmann's account of justification. He defends, however, that his objection also applies to Alvin Plantinga's account of warrant.

> sure you can't see any lakes from the peak?" and gives me a knowing look. I, confused, think back – and sure enough, remember my visual experience of a lake and rather shamefacedly abandon my no-seen lakes belief.) While I am at the peak, I form other beliefs – I see a rock next to me, and form the belief that there is a rock at the top of the peak; I see a tree on the mountainside and form the belief that there is a tree there. […] Clearly, my belief that you cannot see any lakes from the peak is no longer justified after my trip to the peak – it has been defeated by my visual experience of seeing a lake. Equally as clearly, my beliefs that there is a rock and a tree at the peak are not defeated by my visual experience of a lake, and so they remain justified" (2011, 436).

The idea here is that, on the one hand, my belief that you cannot see any lakes from the peak has been defeated by my visual experience of seeing a lake. And, on the other hand, my beliefs that there is a rock and a tree at the peak remain justified. Johnson's contention is that a proper functionalist account of justification or warrant cannot deliver these same results.

The defeater of my belief that you cannot see any lakes from the peak is what Johnson calls a should-be-believed defeater. He takes such defeater to be a subset of normative defeaters. As we have seen, a normative defeater is acquired when one should have taken one of her beliefs to be defeated by another belief or experience but this did not occur. In a proper functionalist account of justification or warrant, properly functioning defeater systems will adequately respond to defeating information. In such an account, a normative defeat is therefore the result of a malfunctioning defeater system. My belief that you cannot see any lakes from the peak would have produced a believed defeater had my defeater system functioned properly.

But Johnson sees a problem with this response. He believes that in cases of defeating experiences the proper functionalist would have to identify or individuate cognitive faculties with the right precision such that a malfunction will defeat what it should defeat. But proper

functionalism seems – or so argues Johnson – to lack the resources necessary to provide this precision in individuation.

In the case above, we find three sets of cognitive faculties in operation:

> (1) The cognitive faculties which produce and sustain the lake-belief.
>
> (2) The cognitive faculties which produce and sustain the rock-and-tree-beliefs.
>
> (3) The cognitive faculties which malfunction in failing to generate a belief that there is a lake and that you can see lakes from the peak.

The proper functionalist must say that set (1), unlike set (2), is malfunctioning, and that the malfunctioning cognitive faculties (3) are the same that produce the lake-belief (1), but different from those of set (2). Thus, the proper functionalist must individuate faculties in such a way that (1) and (3) must count as the same faculties and (2) and (3) must count as different faculties. And Johnson defends that this is not something that the proper functionalist can do. (1) is acquired via testimony and sustained via memory, (2) is acquired via vision, and the malfunctioning faculties (3) are also visual faculties. (2) is exercised and (3) fails to be exercised while I am at the peak. So the proper functionalist will not be able to individuate faculties in such a way that (1) and (3) count as the same faculties and (2) and (3) as different faculties. If she individuates (3) in such a way that it will be the same as (1), it will also count as the same as (2). And if (3) is individuated in such a way that it will count as different than (2), it will also count as different than (1).

Thus, defends Johnson, the impossibility, given proper functionalism, of individuating cognitive faculties with the right precision so that malfunctions will defeat what it should defeat, in this case my no-seen-lakes belief. The heart of Johnson's argument, as he summarizes it, is that

> We cannot account for the fact that the visual experience of a lake is relevant for the justification of the no-seen-lakes belief but not my other beliefs solely in terms of proper function. There is no greater distance in terms of cognitive faculties and processes between the visual experience of the lake and the rock-and-tree-

beliefs than there is between the visual experience of the lake and the no-seen-lakes belief. However, you individuate cognitive faculties and processes, you aren't going to be able to connect the visual experience of the lake and the no-seen-lakes belief without also connecting it to a bunch of other beliefs that it shouldn't be justificationally relevant for (2011, 439).

THE CHALLENGE OF ANTI-DEFEATISM

The idea that justification and knowledge are defeasible has become one of the cornerstones of contemporary epistemology. As a result, most theories of justification or knowledge have sought to make room for the notion of defeat. Recently, however, several challenges to this notion have emerged. In this chapter, I present some of the alleged problems with the notion of defeat in epistemology as developed by Maria Lasonen-Aarnio and Max Baker-Hitch and Mathew Benton. I conclude by suggesting that two of the evidentialist theories of epistemic support that we examined in the first part of this essay can potentially escape those problems, which increases the plausibility of these accounts of support in comparison with the externalist and internalist alternatives that face the problems raised by Lasonen-Aarnio and Baker-Hitch and Benton.

Introduction

The concept of epistemic defeat, or some surrogate for it, is," as Jonathan Kvanvig reminds us, "essential for any fallibilistic "epistemology" (2007, 107). The notion of defeat, at least as it has been conceived in Gettier's wake, has come, nonetheless, under heavy fire recently. Some (such as Bayesians) have found it very difficult to harmonize defeat with their favored formal apparatus.[101] Others (such as Maria Lasonen-Aarnio) consider the existence of a defeater for a certain belief irrelevant for whether that belief constitutes knowledge. Still others find it difficult to implement the notion of defeat in epistemology (Max Baker-Hitch and Mathew Benton). In what follows I present the criticisms developed by Maria Lasonen-Aarnio and Max Baker-Hitch and Mathew Benton.

LASONEN-AARNIO'S OBJECTION

[101] Discussion of such or related problems can be found in, for instance: Kotzen, Matthew (2010) A Formal Account of Epistemic Defeat; Pryor, Jim (manuscript) Uncertainty and Undermining; Weatherson, Brian (2007) The Bayesian and the Dogmatist. Proceedings of the Aristotelian Society 107: 169-85; Wiesberg, Jonathan (2009) Commutativity or Holism? A Dilemma for Conditionalizers, British Journal for the Philosophy of Science; Wiesberg, Jonathan (2015) Updating, Undermining, and Independence, British Journal for the Philosophy of Science.

Maria Lasonen-Aarnio (2010) has challenged the widely accepted assumption that knowledge is defeated in cases of misleading evidence. She makes her case for this view through two cases:

> *Trick on Suzy*
>
> At a time t_1 Suzy comes to know that a certain object is red based on perception. There is nothing abnormal about her perceptual abilities or the lighting in the room. At a slightly later time t_2 a highly reliable and trustworthy authority tells her that the object is illuminated by peculiar red lighting, lighting that would make objects of any color look red.
>
> *Fred's draws*
>
> Fred places exactly one red and one black ball into a bag. Based on perception, at a time t_1 he knows that there is a black ball in the bag. He then starts making draws with replacement, carefully observing the outcome of each draw. Throughout, he remains certain that the contents of the bag don't change. By a later time t_2 he has made ten thousand draws, each of which has produced a red ball (2010, 1).

In both cases, the subject acquires a rationality defeater through misleading evidence. Suzy acquires misleading evidence against her belief that she is seeing a red object when she receives reliable testimony to the contrary. Fred acquires misleading evidence against his belief that there is a black ball in the bag when, after ten thousand draws, he fails to draw a single black ball. According to the dominant view among epistemologists (shared by both internalists and externalists), in both cases the subject has lost knowledge by losing justification for believing the target proposition, even if they continue to believe the target proposition after acquiring the misleading evidence.

Lasonen-Aarnio contends that this seems to pose a problem in particular for externalist theories of knowledge ("Why," she asks, "can't Suzy's and Fred's beliefs continue to be held by a perfectly reliable method?" (2010, 2).[102, 103] To show this Lasonen-Aarnio takes

[102] She says that some externalist attempts, such as Goldman's, at incorporating the notion of defeat into their theories strike her as "being dangerously close to dealing with defeat by brute force" (note 7, pp. 18-9).

knowledge to be safe belief and concludes that she has "found no compelling reason to think that beliefs retained in defeat cases are always unsafe" (7). In Suzy's and Fred's cases, the result the defender that there is knowledge defeat in such cases wants to see is that after acquiring the misleading evidence the subject's belief is no longer safe. But to obtain such a result she will have to say that the method originally employed by Suzy and Fred in the formation of their beliefs in the target proposition is no longer available once they acquire the misleading evidence and that no other method that leads to safe belief in the target proposition is available. As Lasonen-Arnio explains,

> [. . .] because Suzy originally knew *r*, at t_1 it must have been the case that her belief in *r* was safe. Hence, at t_1 she could not easily have falsely believed the object to be red. Now, if Suzy simply retains a belief in *r* on the same basis as before, it is difficult to see how her later belief could fail to be safe. For presumably, it remains true throughout that Suzy could not easily have initially formed a false belief about the colour of the object. Moreover, there is no reason to think that the mere retention of her belief introduces a failure of safety, assuming that Suzy could not easily have, at t_2, ended up believing the relevant object to be any colour other than red (2010, 5).[104]

She defends that the strategy the defender of safety will have to adopt for obtaining the result that Suzy's and Fred's beliefs in the target proposition are no longer safe once they acquire the misleading evidence will have to include the claims that:

---

[103] She also believes that the traditional defeasibility theory has difficulties accounting for knowledge defeat: "Note that such a defeasibility analysis of knowledge is in tension with the idea that knowledge, and not just justification, is defeasible. For if having knowledge is incompatible with the existence of true propositions that would defeat one's justification, then knowledge cannot be defeated, at least not through justification defeat." (Note 4, p. 18).

[104] And the lesson to be derived from this is that:

"if Suzy continues believing *r* on the very same basis as before, then as long as the mechanism by which she retains her belief is such that it could not easily have led her to believe the object to be any colour other than red, her belief in *r* is safe at t_2 as long as her belief in r is safe at the earlier time t_1 – which by assumption it is." (Ibid.).

(1) When the misleading evidence is acquired, the subject will have to cease holding the belief on the same basis as before. "She must rebase her belief by undergoing a token process exemplifying some method of belief-formation" (2010, 5).[105]

(2) The safety theorist will have to defend the idea that "once the new evidence has been acquired, there is no method available to the subject by which she could come to form a safe belief in *p*" (2010, 5).[106]

The general idea that emerges from this attempt by the safety theorist to explain defeat is that "the token processes giving rise to beliefs in defeat cases are always flawed in a way that prevents the resulting beliefs from counting as knowledge" (2010, 5). Lasonen-Arnio explores some possible explanations available for the defender of safety for this failure in the method producing the belief and finds them all unsatisfactory.[107]

---

[105] The distinction between methods and bases is expressed by Lasonen-Arnio as follows:

"Let me first introduce a distinction between believing a proposition on the same basis as before and believing a proposition by employing the same method as before. Forming a belief at least often involves a token application of an epistemic rule or method, but one might think that such a process does not occur at any instant of time during which the belief is retained. I will express this idea by saying that by simply retaining a belief a subject continues to believe p on the same basis as before. This should not be confused with believing p based on the same method as before, since a subject might undergo a new belief-forming process that exemplifies the same method as the one she previously used. Assume, for instance, that Suzy comes to believe that Tuesday is a rainy day based on looking outside in the morning. She then immerses herself in a book, ceasing to consciously entertain the belief, and forgetting all about the rain. At noon she takes a look outside again, and undergoes a new mental process producing a belief that Tuesday is a rainy day. In this case, she may well believe that Tuesday is a rainy day by using the same method as before, but not on the same basis as before." (Ibid., pp. 5, 6).

[106] And she thinks about availability in the following terms: "If a method telling one to believe p in circumstances C is not available to a subject, then for some reason the subject cannot successfully follow that method: she cannot come to believe p as a response to being in circumstances C. This might be the case, first, simply because she is not in the right circumstances and hence, cannot do anything as a response to being in them. But, second, even when in the right sorts of circumstances, she might not have the ability or right know-how to employ the method."

[107] Here are two such explanations: "For the above idea to work, it looks like the method originally employed by Suzy can no longer be available for believing *r* once the defeating evidence is acquired. For if Suzy's original perceptual method is still available, and it produced knowledge of *r* at the earlier time t_1, why would it not still do so? The safety theorist cannot exploit the thought that if a reliable person says that there is trick lighting, then there could easily have been trick lighting, for if this was what prevented the belief formed at t_2 from counting as knowledge, it would also prevent the belief formed at t_1 from counting as knowledge. I argued above that sometimes cases in which a subject doesn't employ a given method will have to be counted as close to cases in which she does. But it's difficult to see how this could help, for there need not be any other

The reason why we are inclined to say there is knowledge defeat in those two cases is, defends Lasonen-Aarnio, because we tend to confuse failure to act reasonably with failure to know. But on the view defended by her, cases of defeat and ordinary cases of knowledge are distinguished by the fact that, when the subject retains belief in defeat cases, she isn't acting in a way that is epistemically reasonable. This is so because reasonable subjects "manage their beliefs in a way that makes sense given the goal of knowledge acquisition," (2010, 12) and belief revision in face of undermining evidence is part of a policy that is knowledge conducive.

The upshot, according to Lasonen-Arnio, is that "subjects who retain beliefs in defeat cases fail to act in an epistemically reasonable manner, but though reasonableness is connected with knowledge in an important way, it is not necessary for knowledge" (2010, 2). Defeaters defeat reasonableness, for sure, but they do not necessarily defeat knowledge. Stubbornly maintaining belief in the target proposition despite the presence of defeat may lead one to continue holding the belief unreasonably but that may have no implication for whether she continues to know the target proposition. "A subject," writes Lasonen-Arnio, "can know despite being unreasonable" (2010, 1: abstract of pre-publication draft). The reason why we are inclined to think that one cannot act unreasonably and yet come to know is because the paradigmatic cases of knowledge are those in which the subject forms the belief in a reasonable way. But, as suggested by Lasonen-Aarnio, there are also (deviant) cases in which the subject knows despite holding the belief in an unreasonable manner, such as in the two defeat cases above. Although "[u]nreasonable subjects are genuinely criticisable, [. . .] like almost anything, knowledge can sometimes be achieved in the absence of a good general strategy" (2010, 1: abstract of pre-publication draft).

BAKER-HITCH AND BENTON'S OBJECTION

---

candidate method that Suzy could easily have employed, a method that might have led her to a relevantly similar false belief. Another thought is that now the fact that misleading evidence is present enters into determining relevant similarity of beliefs or belief episodes: the question to ask is whether Suzy could easily have formed a false belief based on perception in the presence of relevant misleading evidence. But this seems to be conceding that the presence of misleading evidence is relevant for how the subject's belief was formed. It looks like a variant of a view on which the original method is not available, since now the presence of misleading evidence is characteristic of the method employed. I now turn to this idea." [Ibid., p. 6]

As mentioned in the previous chapter, Max Baker-Hitch and Mathew Benton take their criticisms of what they call "defeatism" to supplement that advanced by Lasonen-Arnio. They propose to go beyond Lasonen-Arnio's contribution "by sketching an account based on the knowledge norm of what makes belief 'unreasonable' in some of cases of putative defeat" (2015, 40). They defend the view that both internalists and externalists have difficulty accommodating the notion of defeat into their favored theories.

The difficulty for internalists comes from their conception of defeat in terms of probability-lowering, *i.e.*, they tend to see defeaters as evidence that is added to the believer's doxastic system in such a way that it reduces one's evidence in support of belief *p* to the point that such belief is no longer justified. Suppose, similarly to what we find in Suzy's case above, that you are looking at a red-looking wall and form the belief (which constitutes knowledge) that "that's a fine red wall!" But then someone comes and tells you that the wall is being irradiated by red light. According to the predominant view of defeat, the evidence that had made it probable that the wall is red has been, as Baker-Hitch and Benton put it, "screened off" and you thus no longer know that proposition you believed.

There are two ways to understand what the defetee in such cases of defeat is. It can be understood as being the proposition that "it *appears* that that's a fine red wall!" or that "that *is* a fine red wall!" If the latter,[108] since you know that the wall is red, the new information about the red light irradiation will not reduce the probability on your evidence that that fine wall is red. If the former, the probability that the wall is red – Pr(*wr*) – does seem to be raised by the appearance that it is red – Pr(*Awr*):

Pr(*wr*/*Awr*) > Pr(*wr*)

And once the defeater (it appears that there is a red light – *Arl*) enters your noetic system, the conditional probability reverts back to the prior probability of the wall being red:

[108] Such views include Timothy Williamson's formula E=K, which, as discussed in chapter 1, says that one's evidence is what one knows, but also "any view that allows more than just appearance propositions into one's evidence will allow that there can be cases where a subject's evidence includes a non-appearance proposition *p* and where the subject subsequently comes to believe a proposition *d* according to which her belief that *p* was formed in a defective way" (Backer-Hitch and Benton, 2015, 42).

Pr(*wr*/*Awr* & *Arl*) = Pr(w*r*)

But the problem, they claim, is that, formally speaking, *Awr* does not favor *wr* over a corresponding skeptical hypothesis (¬*wr* & *Awr*), and may even favor the latter. Baker-Hitch and Benton explore the different results that can be obtained in assessing which hypothesis will be favored by conditionalization on *Awr* – *wr* or the skeptical hypothesis – using the Law of Likelihood[109] or the Weak Law of Likelihood,[110] and taking one's prior for Pr(*Awr*/*wr*) to be either <1 or =1. Since (¬*wr* & *Awr*) entails *Awr*, the appearance will either confirm the skeptical hypothesis[111] or be evidentially inert. Thus, any attempt to model defeat on a Bayesian framework with knowledge being gained by updating on appearances, will face a serious problem.

Externalists have usually included a no-defeater condition on their account of knowledge. Baker-Hitch and Benton see this, however, as producing a "gerrymandered picture of knowledge (or justification)," (2015, 45)[112, 113] with no connection with what characterizes the externalist view about knowledge. They explore ways in which purely externalist accounts of defeat could be constructed and find all of them unsatisfactory.

First, they explore method-switching, in both internalist and externalist varieties of method-individuation. The general idea is that defeat occurs when, in face of new evidence, there is

[109] A piece of evidence *E* favours a hypothesis H_1 over a competing hypothesis H_2 just in case: Pr(E/H_1) > Pr(E/H_2)

[110] *E* favours H_1 over H_2 if: Pr(E/H1) > Pr(E/H_2) and Pr($E/\neg H_1$) ≤ Pr($E/\neg H_2$)

[111] In what may have been the original version of this objection, Roger White (2006) pointed out that, given the Law of Likelihood, and given that Pr(*Awr*/¬*wr* & *Awr*) = 1, we have Pr(*Awr*/¬*wr* & *Awr*) > Pr(*Awr*/*wr*), i.e., *Awr* will favor the skeptical hypothesis (¬*wr* & *Awr*) over *wr*.

[112] See footnote 104 above for Lasonen-Aarnio's views on how certain externalist attempts at incorporating defeat into their picture of knowledge seem to come close to amounting to "brute force."

[113] As they put it, "For externalists about justification, however, simply adducing an internalist no-defeat condition on knowledge (or justification) is not an attractive option. Doing so results in a gerrymandered picture of knowledge (or justification) that includes an internalist defeat condition bearing no connection with the deep structural feature that they, qua externalists, take to be characteristic of knowledge (e.g. safety, sensitivity, track-record reliability, aptness)" (Ibid.).

change in the method that produces the belief such that the new method is unreliable.[114] Whereas the externalist view of method individuation takes methods to be individuated "in terms of actual facts about how the subject's belief is sustained" (2015, 46) internalist method individuation consists in how it seems, from the perspective of the subject, that her beliefs are formed. Baker-Hitch and Benton believe, that both the internalist and the externalist varieties fail to provide an adequate account of defeat, for – following Lasonen-Arnio – "it is simply not plausible that in all defeat cases the subject counts as having her belief sustained by a method that is unreliable" (2015, 46).

Secondly, they explore defeating evidence as a determinant of which cases are close when knowledge is conceived in terms of certain formulations of the safety condition of knowledge, such as Timothy Williamson's.[115] Here defeat is construed in terms of new information that alters which cases are counted as close. If closeness is taken to be determined by a subject's experience then perhaps, as John Hawthorne has suggested, it may be possible to accommodate defeat into a safety framework.[116]

Baker-Hytch and Benton find two main problems with such a proposal. First, the prioritization of an experience-based account of similarity over other accounts – such as the etiological one – can have skeptical results. Second, this account may not be capable of ruling out as cases of defeat those that involve obviously false information – such as you being told that you are not cold when you know that you are feeling cold.

---

[114] This could be tried with any externalist theory that makes room for method-individuation, such as the truth-ratio variety of reliability, as well as modal principles of knowledge such as sensitivity and safety.

[115] Williamson's safety can be formulated as follows: If one knows, one could not easily have been wrong in a similar case (2000, 147).

[116] Here is how Hawthorne's suggestion goes:

"Consider a case in which you competently perform a calculation using a fully working calculator and get the correct answer. Plausibly there are no cases of error that are sufficiently close to this case in the respects mentioned earlier, and so your belief counts as safe. Suppose, however, that some time later Devious Dave tells you (albeit entirely misleadingly) that the calculator has a wiring defect such that it churns out mistaken answers half of the time. In line with Hawthorne's suggestion, the fact that you undergo this experience involving Dave's testimony makes a difference to which cases count as close. In particular, other cases in which you undergo the same experience (the one involving Dave's testimony) will now count as sufficiently close to the target case, and some of those other cases, so the thought goes, will be ones in which the calculator's wiring really is faulty so that you end up believing falsehoods. Hence, your belief is no longer safe" (2015, 50).

Thirdly, they evaluate the prospects for an alternative reliable processes approach to defeat. According to this account, originally developed by Alvin Goldman,[117]

> *S*'s belief that *p* is defeated if (and only if) *S* has available to her an alternative reliable process which is such that had she employed that process in addition to the one she actually used, then she would not have continued to believe that *p* (2015, 51).

Suppose you are told that the computer lab of building 5 doesn't open on Saturday, but later on somebody else tells you the person who gave you this information is mistaken – the lab does open on Saturday morning. On Goldman's account, your belief that the lab doesn't open on Saturday is defeated by the second testimony. And the reason for that is because there is an alternative reliable process such that had you formed a belief through it you would have ceased to believe that the lab won't be open on Saturday.[118]

But Baker-Hytch and Benton believe they have found two flaws – in addition to counterexamples developed by Bob Beddor[119] – in Goldman's account of defeat. First, to take the case under discussion, for Goldman's account to work, token processes would have to be assigned to narrower types than simply testimony, *i.e.*, testimonial process types would have to be individuated. This, however, brings a host of additional problems characteristic of method individuation. Second, Goldman's account seems to deliver counter-intuitive results in certain types of cases that suggest that his account lacks "the

[117] Goldman, Alvin (1979) What is Justified Belief?

[118] It's important to note that, on Goldman's account, the alternative process must be available to the subject. You could have decided to consult a third friend about the business hours of the lab. That process, unlike the first two testimonies, which are available to you through memory, would not be available to you. As Goldman put it,

"[I]t seems implausible to say all "available" processes ought to be used, at least if we include such processes as gathering new evidence. Surely a belief can sometimes be justified even if additional evidence-gathering would yield a different doxastic attitude. What I think we should have in mind here are such additional processes as calling previously acquired evidence to mind, assessing the implications of that evidence, etc." (cited in Baker-Hytch and Benton, 2015, 51).

[119] Beddor, Bob (2014) Process Reliabilism's Troubles with Defeat.

normative universalizability that an account of defeat should be able to deliver up" (2015, 53).

They conclude their criticism of externalist views about defeat by assessing the idea that defeat cases should be seen as cases in which properly functioning noetic systems respond to defeating evidence with suspension of belief and by exploring the prospects for a virtue-theoretic account of defeat. Since we discussed their criticism of the proper functionalist account of defeat in the previous chapter, let us conclude this summary of their objections by mentioning that, on their view, the challenge faced by the virtue theorist is to explain how a defeater can turn an apt performance into an inapt one. And, given the criticisms of the externalist views that we have seen so far, it seems fair to say that the task of the virtue theorist seems particularly daunting, for, as the authors put it,

> gaining evidence that your performance was inapt is not itself enough to make your performance inapt; likewise gaining evidence that you don't know may not itself be enough to make one not know (2015, 55).

As mentioned above, their second project in the article is to provide an alternative – and uniquely successful – account of defeat in terms of knowledge-centered epistemic rationality. They postulate that belief is governed by the following epistemic norm (KNB) and its derivative (KNBa):

> (KNB) One must: not believe that *p* if one does not know that *p*.
>
> (KNBa) One must: refrain from believing *p* if one comes to believe or accept that one's belief that *p* is not knowledge.

In possession of these two principles, they defend that cases of defeat are cases of violation of KNBa. Cases of defeat, they claim, are cases in which it would be irrational or irresponsible for the subject to continue to maintain her belief. Like Lasonen-Arnio, they defend that subjects who continue to believe in defeat cases are genuinely criticizable. But while Lasonen-Arnio defends that subjects in such cases are criticizable for not following policies

that are rational for those seeking to acquire knowledge, they see these subjects as violating the derivative norm of belief (KNBa). Thus, when one acquires an undercutting defeater (*i.e.*, evidence that one's belief that *p* is not conducive to the truth that *p*), what happens, given this picture of defeasibility, is that, if one continues holding the belief, she will be violating KNBa. In other words, cases of undercutting defeat are cases in which one gains indication that if she continues to hold a certain belief she might violate KNBa and thus that the correct response is for her to give up the belief.

Their account of defeat deals with rebutting defeaters (i.e., evidence not only of a lack of connection to the truth of the belief, but that the belief is false) in much the same way. Since in such cases one acquires evidence that the belief is false, so here, too, the subject acquires evidence that her belief is not knowledge. And, as a result, non-violation of KNBa would require that one gives up the belief. But note that, because one can be wrong about whether one knows, violation of KNBa may not incur in violation of KNB itself. And this, they claim, is consistent with beliefs in certain defeat cases, such as those advanced by Lasonen-Arnio, continuing to constitute knowledge.

In any case, for both undermining and rebutting defeaters, they claim, the probabilistic model of defeat will not do the job. The problem, according to them, is that

> in particular, there is no uniform way of spelling out what it is for a defeater to be added to one's evidence which enables the defeatist to have both of what they want, namely for the defeater to lower one's probability for some *p*, but for the evidence (apart from that defeater) to confirm *p* over skeptical hypotheses concerning *p* (2015, 60).

But the externalists, they claimed, fare no better. Each of the major externalist theories of justification and knowledge would have serious problems in attempting to formulate a satisfactory account of defeat. Alternatively, they offered a new approach to epistemic defeasibility in line with E=K that, as they see it, "allows that one can retain knowledge in the face of cases of putative defeat" (2015, 60).

ALTERNATIVE ACCOUNTS OF DEFEAT

In their survey of possible formulations of epistemic defeat, Baker-Hitch and Benton did not discuss two evidentialist alternatives to the Bayesian account of epistemic support that we explored in the first part of this essay, namely, dispositionalism and explanationism. In the remainder of this chapter, we will briefly examine whether accounts of defeat formulated along dispositionalist and explanationist lines can give more satisfactory results than those discussed above.

First, then, how could a dispositionalist account of defeat be formulated in a way that avoids the problems that affect internalist versions of defeat that rely on Bayesian principles of epistemic support? Recall that, on Byerly's dispositionalism (2014), *S* is justified in believing *p* when she is disposed, in light of her total evidence, to believe *p*; *S* is justified in disbelieving *p* when she is disposed, in light of her total evidence, to disbelieve *p*; and *S* is justified in suspending judgement about *p* when she is disposed, in light of her total evidence, to suspend judgement about *p*. And recall that, as described above, Suzy and Fred cease to be justified in believing the target propositions after they acquire defeaters through misleading evidence. Suzy acquires misleading evidence against her belief that she is seeing a red object when she receives reliable testimony to the contrary. And Fred acquires misleading evidence against his belief that there is a black ball in the bag when, after ten thousand draws, he fails to draw a single black ball. According to the dominant view among epistemologists, in both cases the subjects have lost knowledge by losing justification for believing the target proposition, even if they continue to believe the target proposition after acquiring the misleading evidence. Lasonen-Aarnio, however, doesn't find any good reason to think that, particularly on externalist approaches, knowledge is lost if the subjects continue to believe the target proposition. And Baker-Hitch and Benton attempt to show that internalist theories that rely on Bayesian versions of epistemic support cannot model defeat adequately.

Dispositionalism, however, as developed by Byerly, is not intended to be a theory of knowledge. So the relevant question here is whether it can circumvent the problems that Baker-Hitch and Benton pose for internalist theories of defeat. When Suzy acquires

misleading evidence against her belief that she is seeing a red object and Fred acquires misleading evidence against his belief that there is a black ball in the bag, their total evidence (TE) changes from TE_1 to TE_2. On dispositionalism, Suzy would lose justification for believing that the object is red if, on TE_2, she loses the disposition to believe that the object is red and develops the disposition to suspends judgment about such proposition; and Fred would lose justification for believing that the object is red if, on total evidence$_2$, she loses the disposition to believe that there is a black ball in the bag and develops the disposition to suspends judgment about such proposition. Hence, the dispositionalist answer is straightforward and doesn't seem to be affected by the problems raised for the Bayesian view. Of course, one could be very idiosyncratic in the way one develops dispositions to believe. Byerly mentions the case of the detective who, half-way through his examination of the evidence available, becomes disposed to believe that *E* is the best explanation of her evidence. Similarly, if the detective had a number of equally good explanations, but was disposed to believe one of them, she would be justified in doing so. On dispositionalism, again, one should believe what one is disposed to believe in light of her total evidence. But that is not the end of the story: "My view," Byerly acknowledges, "is not that a detective who is in fact so disposed ought not believe *E*, but rather that a detective who is so disposed has no business being a detective"[120] (407-408). In order to escape this untoward consequence, Byerly adds a virtue component to his account of justification that includes a condition of epistemic value explained in terms of "a believer's performing her proper function and doing so with excellence" (420). This synthesis of dispositionalism and virtue epistemology is expressed by the following condition of justification:[121]

> (VirtPFE) A person *S* fulfills her proper function as a believer with excellence to the extent that she takes all and only those doxastic attitudes which she is sufficiently strongly disposed to take by *virtuous dispositions* in light of all of her evidence (2014, 421).

[120] And Byerly notes that Feldman has offered similar responses to similar objections (see 2014, 408, footnote 25).

[121] So, supposedly, Byerly is committed to two views of rationality, one internal (in terms of developing the right sort of dispositions) and one external (in terms of developing the right sort of virtues).

Thus, while the detective is justified in believing flunky propositions because she is *strongly*[122] disposed to do so, she is not a good detective, one who carries her duties with epistemic excellence, fulfills her proper function as a believer and detective.

But recall that the problem raised by Baker-Hitch and Benton for the Bayesian account was that, on the Bayesian way to formalize undermining defeat, we get the result that *p* is not favored over the skeptical hypothesis. Could a similar problem be raised for dispositionalism? I don't think so. For, while Bayesianism is concerned with the conditionalization of evidence under the constraints of the axioms of probability theory, dispositionalism only tells us that rationality is a matter of how one is disposed to believe in light of her total evidence. Dispositionalism is construed by Byerly as a theory capable of circumventing what he perceives to be the overintellectualization of epistemic support by other theories. In fact, I suspect some may find dispositionalism to be too alike externalist theories of knowledge and justification, and this may be particularly the case when it comes to a dispositionalist response to the skeptical problem – instead of facing it head on, the dispositionalist is likely to simply say that the fact that one doesn't have a good response to skepticism doesn't mean one does not have non-inferential justification and knowledge, for justification (and knowledge), on this account, doesn't rest on one having direct access with the facts.[123]

Unlike the externalist (and perhaps the dispositionalist) view of the skeptical problem, the explanationist view is that the skeptic should and can be faced head on. Explanationists such as McCain (2014), Vogel (1990), Moser (1989), and Lycan (1988), have defended responses to skepticism that take the commonsense hypothesis to be capable of explaining certain features of our sensory experience of the world better than the skeptical hypothesis. McCain (2014, chapter 6.3), for instance, comparatively evaluates these two hypotheses in terms of four important explanatory virtues (quantitative parsimony, qualitative parsimony, explanatory simplicity, and explanatory questions) and argues that the commonsense hypothesis performs better on two of them (the last two), with the remaining two delivering

[122] The "strongly" condition is added to differentiate dispositionalism from epistemic conservatism (2014, 417).

[123] See Poston, Ted (2008) for this way of formulating the externalist response to skepticism.

inconclusive results. Explanationism seems, then, well positioned to deliver an account of defeat that can potentially satisfy the concerns expressed by Baker-Hitch and Benton. For, if, on explanationism, it is plausible that defeaters perform their defeating role by being the best explanation for certain facts in light of our total evidence, and that explanatory considerations favor the commonsense hypothesis over the skeptical one, then it seems plausible that, in at least many cases, the defeater, as the best explanation for certain facts, will also rule out the specter of skepticism.

On McCain's explanationism, one is justified in believing the best available explanation of one's mental states. Availability is understood in terms of dispositions to have a certain sort of seemings. One doesn't need to have concepts of "explanation" or "evidence," but merely of "'the concepts required to understand *p*' and be disposed to have a seeming that *p* is part of the best answer to the question 'why does *S* have *e*?'" (2018, 10). Thus, when Suzy acquires misleading evidence against her belief that she is seeing a red object when she receives reliable testimony to the contrary, her belief about the object ceases to be part of the best explanation for her experience of seeing a red object. Her total evidence now includes the experience of the reliable testimony to the effect that the object is being illuminated by red lighting. As a result, we should expect that she will cease to be disposed to have a seeming that the proposition about there being a red object in the room is part of the best explanation for her experience of seeing a red object. Likewise, when, after ten thousand draws, Fred acquires misleading evidence against his belief that there is a black ball in the bag, his belief about there being a black ball in the bag ceases to be part of the best explanation for his experience of putting a black ball in the bag. His total evidence now includes the experience of witnessing ten thousand draws without seeing a single black ball emerging from the bag. As a result, it is to be expected that he will cease to be disposed to have a seeming that the proposition about there being a black ball in the bag is part of the best explanation for his memory of placing a black ball in the bag. Thus, explanationism seems well-positioned to provide an account of epistemic defeat that is immune from the criticisms leveled by Baker-Hitch and Benton against Bayesian accounts of defeat. In fact, explanationism apparently has significant virtues when it comes to epistemic defeat. According to McCain, explanationism

handles both rebutting and undercutting defeaters in a non-ad hoc manner. Defeat is simply a natural extension of the core idea that the best available explanation of S's total evidence determines what she has justification for believing. Another potential advantage of Ex-PC[124] when it comes to defeat is in terms of simplicity. Ex-PC offers a picture of justification that can do away with talk of epistemic defeat altogether while still getting the intuitively correct results in various cases where philosophers typically make reference to defeaters. It can do this without simply building in a "no-defeaters" clause. One might think that this marks a significant advantage of Ex-PC because it offers a much simpler and more unified picture of justification than what is offered when a theory of justification has to add additional features to accommodate the work of defeaters (2018, 12).

[124] Ex-PC refers to McCain's combination of explanationism and phenomenal conservatism that he proposes and develops in his 2018 paper.

COGNITIVE SCIENCE OF RELIGION AND THE RATIONALITY OF THEISTIC BELIEFS

In this chapter, I review of the contemporary discussion on the implications of the findings of Cognitive Science of Religion (CSR) for the rationality of theistic beliefs. We will explore the defeating potential for the rationality of theistic beliefs of twelve objections found in the literature: The Natural Explanation Objection I, the Natural Explanation Objection II, the Neural Substrate Objection, the By-product Objection, the Religious Utility Objection, the Inherited Beliefs Objection, the Lack of Proper Causal Relationship Objection, the False Positives Objection, the Mutually Exclusive Beliefs Objection, the Simplicity Objection, the Problem of Natural Non-Belief, and the Confirmation Bias Objection. I have attempted to order the presentation of the objections according to their strength – from the weakest to the strongest, or at least to the most sophisticated ones or to those that have generated the most interesting and promising discussions (no doubt with a certain degree of subjectivity in this selection).

BACK TO RATIONALITY AND DEFEATERS

As we will see below, most CSR objections to theistic beliefs come in the form of a charge of unreliability of the mechanisms or faculties that produce such beliefs. The link between reliability and rationality, as we have seen, belongs ordinarily to the domain of external rationality, *i.e.*, rationality with respect to the proper function or reliability of one's cognitive faculties, independently of one's awareness of the factors that contribute to their reliability. Recall that one's beliefs can be internally rational even when they are externally irrational, and internally irrational even when one's beliefs are externally rational. While internalists and evidentialists are concerned with internal rationality, *i.e.*, with the formation of beliefs in response to experience, if one acquires evidence that the deliverances of one of her cognitive faculties is not reliable, she will, in the absence of defeater deflectors, acquire a defeater for her justification to believe that her cognitive faculty is reliable and, hence, that the outputs of such a faculty are true. That is, her total evidence will indicate that beliefs formed via that specific mechanism should not be trusted and, hence, the rational response in this case is to suspend judgment about the truth-value of beliefs delivered by such faculty.

One's propositional justification/rationality and defeat are a matter of one's total evidence. On evidentialism, one has propositional justification to believe *p* when, roughly, her total evidence supports *p*. And on mentalist evidentialism, in particular, one's total evidence will be constituted by the totality of her mental states. But perhaps there is counterevidence *e* outside of her mental life that would prevent her from believing *p* rationally were *e* to become part of her total evidence. This counterevidence can be misleading or genuine: it can mislead the subject into believing that something that is actually the case is not the case; or it can put one in contact with how things really are. Let us see a case by Lehrer and Paxson (1968) that will help us see more clearly the distinction between misleading and genuine defeat:

> Suppose I see a man walk into the library and remove a book from the library by concealing it beneath his coat. Since I am sure the man is Tom Grabit, whom I have often seen before when he attended my classes, I report that I know that Tom Grabit has removed the book. However, suppose further that Mrs. Grabit, the mother of Tom, has averred that on the day in question Tom was not in the library, indeed, was thousands of miles away, and that Tom's identical twin brother, John Grabit, was in the library. Imagine, moreover, that I am entirely ignorant of the fact that Mrs. Grabit has said these things. The statement that she has said these things would defeat any justification I have for believing that Tom Grabit removed the book, according to our present definition of defeasibility. Thus, I could not be said to have nonbasic knowledge that Tom Grabit removed the book.
>
> The preceding might seem acceptable until we finish the story by adding that Mrs. Grabit is a compulsive and pathological liar, that John Grabit is a fiction of her demented mind, and that Tom Grabit took the book as I believed (Lehrer and Paxson, 1969, 228).

Suppose, however, in a modified description of the case, that, after I saw Tom grab the book, it was brought to my attention that Mrs. Grabit stated that Tom was thousands of miles

away, and that it was John, Tom's identical twin brother, who was in the library. In this case, my belief that Mrs. Grabit said these things would provide me with a misleading (Mrs. Grabit lied about Tom having a twin brother) defeater for my justification to believe that Tom grabbed the book. Lehrer and Paxson's goal in formulating this case was to show that analyses of knowledge along the lines defended by proponents of the defeasibility theory of knowledge up to that moment were unsuccessful due to their failure to make a distinction between genuine and misleading propositional defeaters. In Lehrer and Paxson's original formulation, both the proposition that Mrs. Grabit said that it was John, not Tom, who was in the library, and the proposition that Mrs. Grabit is a compulsive liar are outside of my mental life and constitute, respectively, misleading and genuine information that affect my propositional justification. In our modified case, we have a misleading mental state defeater (Mrs. Grabit's statement that it was John, not Tom, who was in the library), and a genuine information that defeats the misleading defeater, which is outside of my mental life, but which, were it to become part of my total evidence, would restore my justification to believe the Tom stole the book.

With this distinction in mind, the question before us is not simply whether the objections that we will present below can defeat the propositional justification of theistic beliefs. Whether these objections can constitute defeaters for one's rationality is a matter of the subject's total evidence. The objections can, in principle, defeat one's rationality for theistic beliefs even if they are misleading or false (if we assume, of course, as most evidentialists do, that evidence can be false), as long as the subject does not have a deflector of the potential defeater in her total evidence. And if defeat occurs, the justification for her theistic belief will remain defeated until she acquires a defeater-defeater for the original defeater, perhaps in the form of one of the responses that we will see below to the objections against the rationality of theistic belief. But what is more interesting for our purposes in this second part of this essay is whether the philosophical and scientific evidence available to us point to the existence of genuine defeaters for the rationality of theistic belief, *i.e.*, whether the objections are such that they are sound and have remained undefeated in the philosophical and scientific debates. Whether an objection in fact constitutes a genuine defeater is a difficult, if not impossible question to answer, for, since we are not omniscient, there may

always be responses to that objection that we are not aware of. Still, the refutation of false and fallacious arguments and objections and the pursuit of the true answers to the philosophical questions before us are the goals to which philosophical inquiry aspires, however elusive they may be in many cases.

So, do the objections bellow and the responses that have been offered to such objections suggest that we have something like one or more undefeated, genuine defeaters for the rationality of theistic beliefs? Or can the responses that we will explore after each objection constitute proper defeater deflectors or defeaters of the original defeaters, *i.e.*, defeater-defeaters that will preserve or restore the rationality of theistic beliefs (assuming they were rational in the first place)?[125]

OBJECTIONS AGAINST THE RATIONALITY OF THEISTIC BELIEFS

Objection 1. *The Natural Explanation Objection I*

According to the first objection, if we can provide natural or scientific explanations of religious beliefs, then there is something epistemically problematic with such beliefs. Michael Alper, author of *The God Part of the Brain* claims, for instance, that "if belief in God is produced by a genetically inherited trait . . . this would imply that there is no actual spiritual reality, no God or gods, no soul, or afterlife."[126] The claim here is that because there are natural explanations of religious beliefs, these beliefs are false. More modest, and tenable, versions of the objection can be formulated in terms of irrationality, however: Once we see that religious beliefs can be explained in scientific or naturalistic terms, one ceases to be justified in thinking that there actually is an object to which those beliefs refer. Thus, in light of the existence of natural or scientific explanations for theistic beliefs, such beliefs can no longer be held rationally.

Are There Plausible Responses to this Objection?

[125] I will leave to reader the task of thinking whether the notions of neutralizing defeater-defeaters, intrinsic defeater-defeaters, proper-function-rationality defeaters, Humean defeaters, purely alethic rationality defeaters, and so on, apply to the debate dialectic that emerges from the objections.

[126] Quoted in Clark and Barrett (2011, 655).

(A) By explaining something, we say nothing about whether what is being explained is true or whether belief in it is justified. As Barrett put it,

> If cognitive neuroscientists manage some day fully to explain the brain activity and evolutionary history of those brain functions responsible for people believing that seventeen times eleven equals 187, seventeen times eleven would still equal 187. Similarly, a complete scientific explanation for why humans nearly universally believe that other people have minds would not suddenly count against whether humans should believe that others have minds (2009, 96).

(B) Theists have been able to present numerous accounts of how God can use natural processes to produce theistic beliefs. There doesn't seem to be any incompatibility or inconsistency between the existence of God and there being natural explanations of belief in God. The fact that there is a natural explanation does not preclude there being also a supernatural one. Both explanations may be true. As Alvin Plantinga put it,

> To show that there are natural processes that produce religious belief does nothing, so far, to discredit it; perhaps God designed us in such a way that it is by virtue of those processes that we come to have knowledge of him (2000).

(C) God doesn't need to be the immediate cause of theistic beliefs for these beliefs to be justified. By guiding the evolutionary process that formed our cognitive systems that produce these beliefs, He is, the theist may claim, their ultimate cause. So, as Clark and Barrett put it,

> learning that the immediate cause of God beliefs involves natural faculties would not show that our God beliefs were untenable after all. In order to show that, the CSR objector would have to

> show that God was not the ultimate cause of our God beliefs. And that they simply have not done (2011, 660).

Objection 2. *The Natural Explanation Objection II*

A more plausible version of the natural objection explanation objection can be formulated, however, in terms of insensitivity of the theistic beliefs to the truth of theism. For, as Joshua Thurrow (2014b, 196) formulates the objection, if we can explain theistic beliefs solely by natural processes, this means that we would hold theistic beliefs even if theism was false. Thus, the processes leading to theistic beliefs are not sensitive to the truth of theism. And since insensitive belief-formation processes are unreliable, theistic beliefs are unreliable. An example given by Thurrow to illustrate this claim is of someone who perceives orange things as red. When she sees an object that looks red to her, she may well be perceiving an object that is actually orange. Should she then trust her perceptual faculties when she identifies something that seems red to her? No, when it comes to her perceptual abilities with respect to red objects, her perception is unreliable and she should suspend judgment about the color of the object in such cases. Similarly, if theistic beliefs are not sensitive to truth and, thus, unreliable, we should suspend judgment about the truth of theism.

Are There Plausible Responses to this Objection?

(A) Insensitivity claims can plausibly work only with respect to contingent propositions. For, if a proposition is necessary, there is no scenario in which it could be false. As a result, there is no way to formulate the objection that one would continue to believe the proposition if it were false. Hence, beliefs cannot be claimed to be insensitive to the truth of necessary propositions. For whenever one forms belief in a necessary proposition, the belief will be true.

Theistic propositions are quite plausibly the sort of propositions that, if true, are necessarily true. Insensitive claims about theistic beliefs thus face the problem that, if God exists, He is, most likely a metaphysically necessary being.[127] That is to say that there is no possible world

[127] Richard Swinburne (*e.g.*, 2016) is a notable dissenter here. He is one of the very few theist philosophers who deny that God is a metaphysically necessary being.

in which He would not exist. If that is the case, then whenever one forms the belief that the proposition "God exists" is true is forming a belief in a necessary truth and such belief cannot be false (Penner, 2018, Braddock, 2018).

(B) Even if theistic beliefs formed on the basis of CSR mechanisms were shown to be insensitive to truth, this would not show that theistic beliefs are irrational or that they do not respond to evidence, for theistic beliefs might be formed rationally and in response to evidence via other mechanisms or sources of justified belief. In other words, theistic beliefs formed by traditional kinds of reasons (via testimony, via religious experience, via appearance of design in nature, via experience of miraculous events, via arguments from natural theology, via answered prayers, and so on) would not be undermined by the CSR considerations (Thurrow, 2014b, 195-6).

Objection 3. *The Neural Substrate Objection*

This objection, so named by Justin Barrett (2007), is implicit in much of the neurotheology literature. According to it, the discovery of regions of the brain responsible for religious experience and the possibility of stimulation of experiences akin to religious experiences using electromagnetic field shows that theistic beliefs are not justified. The idea here is that since religious experiences can be induced by purely natural means, one is not justified in claiming a supernatural origin for such experiences.

Are There Plausible Responses to this Objection?

(A) The objection presupposes that neurotheology has indeed identified of regions of the brain responsible for religious experience, which is still a matter of controversy.

(B) More importantly, the objection presupposes that natural processes cannot be caused by the supernatural. It presupposes, as Barrett put it, that "the supernatural does not regularly causally act upon the neural substrate" (2007, 61). But this is an unwarranted presupposition. There doesn't seem to be good reasons to think that God could not or would not use natural processes to produce religious experiences.

(C) Instead of defending that the supernatural causally act upon the neural substrate, the theist could simply maintain that God arranged the natural order in such a way that "human brains naturally give rise to religious experiences under particular situations" (2007, 61).

Objection 4. *The By-product Objection*

An objection that can be derived from a by-product view of CSR data, which holds that religious beliefs did not develop and persist because they conferred selective advantage, but as by-product of cognitive mechanisms designed for other purposes, says that, since religious beliefs are the result of processes that did not confer selective advantage, they are mere accidents or illusions and therefore cannot be rationally believed (Barrett, 2007, 62-63; Clark and Barrett, 2011, 662).

Another way to formulate this objection is, as Michael Murray put it, that "religious beliefs are unjustified because the mechanism that produces them was not properly subject to the winnowing power of natural selection" (2009, 176-7). As he explains,

> If my visual system produces beliefs that are largely incorrect, natural selection will catch up with me. Because of this, I can have some confidence that my visual systems are reliable. But supernatural beliefs generated by HADD and other cognitive tools are not subject to the winnowing power of natural selection in this way (2009, 177).

Are There Plausible Responses to this Objection?

(A) For the theist, the natural world is not all there is, and although she may claim that belief in God is something that may have as a proximate cause the natural mechanisms described by CSR, for her the ultimate cause of such belief is to be found in God. The theist can claim that "God configured evolutionary history to make belief in supernatural reality easy or natural for us" (Murray, 2009, 178).

(B) As Michael Murray (2009, 177) notes, this objection would apply to many of our non-religious beliefs, including beliefs in the reality of the external world and in the existence of

other minds, for they lack connection to fitness, too. Thus, if religious beliefs are not justified because they, on a by-product view, lack connection to fitness, so would our cherished beliefs in the reality of the external world and in the existence of other minds.

(C) In particular, the objection would apply to all scientific beliefs, which would include beliefs such as that the theory of evolution is true and about the findings of CSR. Such beliefs, as Barrett put it,

> arose too late in our history to have played a role in natural selection of humans. Evolution did not select for calculus, quantum theory, or natural selection. Are these beliefs then suspect for being 'accidents' or 'byproducts' of evolution? With this line of reasoning, Darwinism would face the ax alongside theism (2007, 63).

(D) Natural selection selects for behavior, not for the truth-value of beliefs. False beliefs can be as adaptive as true beliefs, and hence, unreliable belief-forming mechanisms can be as adaptive as reliable ones. And as Alvin Plantinga (*e.g.*, 2012) and others (Murray, 2009, 177) have argued, given this picture of natural selection, global skepticism looms for the naturalist who sees the improbability of the reliability of our belief-forming mechanisms on unguided evolution. The only way out of global skepticism, Plantinga argues, seems to be by the adoption of a supernaturalism picture of our evolutionary history, *i.e.*, with God as superintending it in a way that the selected mechanisms would be truth-conducive. So it seems that this objection is more of a problem for the naturalist than for the theist, if a problem at all for the theist.

Objection 5. *The Religious Utility Objection*

While the previous objection assumed a by-product view of CSR findings, the current objection assumes a group selection interpretation of CSR. Such a view takes religious beliefs to have adaptive value because of their contribution to social cohesion. But, according to the objection, since such beliefs are selected not by their truth-value, but by their utility, they are not justified (Barrett, 2007, 64-65).

Are There Plausible Responses to this Objection?

God could have designed the evolutionary process in such a way that, via the mechanisms of group selection, we would come to have cognitive faculties that would make us disposed to have religious beliefs. There doesn't seem to be good reasons to think that such faculties, while arising due to utility, rather than truth, would not be reliable and produce true beliefs. Why couldn't God superintend the evolutionary process so that through these natural mechanisms we would come to have faculties that would give us true beliefs about God?

Objection 6. *The Inherited Beliefs Objection*

As formulated by Barrett (2007, 66):

> People are credulous recipients of theistic beliefs (*e.g.*, from parents). Natural selection provided people with the cognitive faculties that make us credulous recipients. As we now know why people so readily believe in gods, continuing to believe is irrational.

And he adds that, according to the objector, given that the CSR mechanisms makes us credulous recipients of theistic beliefs, we cannot step back, so to speak, from those beliefs and soberly examine the grounds we have for such beliefs. And once we see this, we should give up our theistic beliefs.

Are There Plausible Responses to this Objection?

(A) Barrett (2007, 66) responds by saying that the objection assumes that the credibility of theistic beliefs comes only from the mechanisms described by CSR. It neglects that theists may have other grounds for theistic beliefs, such as arguments from natural theology and other reasons and evidence, that work independently of the mechanisms described by CSR.

The objector may respond that the intuitions operative in the CSR mechanisms also operate in other sources of theistic beliefs (See responses to objection twelve for responses to this consideration), or by formulating a weaker version of the argument, which, according to Barrett, says that the objector can grant that some people do have additional reasons, but

many don't, and that those who don't would lose the only grounds they had for theistic belief.

(B) In response to this reformulated inherited beliefs objection, Barrett (2007, 66) says that whether beliefs formed via the CSR mechanisms are justified or not is not a matter of whether the mechanisms incline us to believe, but, rather, of whether they incline us to form mostly true beliefs or not (see response to objection five for a line of reasoning that can strengthen this response). Only if the latter is the case, claims Barrett, can the objector allege that those mechanisms are defective.

Objection 7. *Mutually Exclusive Beliefs Objection*

This objection, formulated hypothetically by Michael Murray (2009), says that religious diversity suggests that the cognitive mechanisms that produce religious beliefs yield beliefs that are mutually incompatible. This mutual incompatibility suggests that the cognitive mechanisms involved are unreliable. Absent independent justification for those beliefs, they should, consequently, be considered unjustified.

Are There Plausible Responses to this Objection?

In order for religious beliefs formed by the cognitive mechanisms identified by CSR to be unjustified because of the religious diversity we see in the world, it would be necessary to show that the inconsistent beliefs are exclusively or primarily the product of such mechanisms, and not of cultural influences. As Murray (2009, 172) put it,

> HADD tells me there is 'an agent'; my beliefs about what sorts of fauna inhabit these parts lead me to conclude that the agent is a bear or a tiger or the bogeyman. If you conclude that it is a bear and I conclude that it is the bogeyman this does not show HADD to be unreliable, it shows that my mom was wrong about the bogeyman. Likewise, no one doubts that divergent cultural traditions play an enormous role in giving religious concepts their specific contours. If the mutually exclusive aspects of these beliefs creep in from cultural sources this does nothing to undermine the

> reliability of these cognitive tools, it just shows that the cultural traditions are false.

Objection 8. *The Simplicity Objection*

According to this objection, discussed by Clark and Barrett (2011) and Murray (2009), an explanation of religious beliefs that does not include the supernatural is simpler than one that includes God as the ultimate cause of such beliefs and should therefore be preferred to the theistic explanation.

Are There Plausible Responses to this Objection?

(A) Simplicity is only one of the considerations or theoretical virtues that must be taken into consideration in the process of theory choice. Simplicity is, as Murray (2009, 175) put it, only *a ceteris paribus* virtue. Consequently, a more complex theory may be preferable to a simpler one if the more complex one fares better with respect to other categories, such explanatory power, coherence, scope, etc. If simplicity were the only relevant criterion in theory selection, we should not believe that other persons exist, for it is simpler to suppose that other persons are just creations of our minds. As Murray explains:

> it is [. . .] simpler to assume that there are no other minds but my own (and that the appearance that other things have minds is just an illusion) than it is to accept that there are many minds. But we don't accept the 'one mind' over the 'many mind' hypothesis because such a hypothesis doesn't cohere with many other things to which we are committed (2009, 176).

(B) Murray (2009, 175-176) argues that even if naturalism is simpler, it is less explanatorily powerful than theism. For theism can plausibly explain a number of features of reality that naturalism has difficulty explaining: objective morality, why there is something rather than nothing, the fine-tuning of the universe for life, and so on.

(C) As Trent Dougherty and Logan Gage (2015) have argued, on the main accounts of the virtue of simplicity defended today (syntactic, ontological, quantitative, qualitative), theism (or God) comes out as simpler than naturalism. And both theistic and naturalistic explanations must end somewhere. The difference is that "Theism postulates one brute fact, and everything else follows from it." Naturalism, however,

> lacks this kind of fundamental explanatory simplicity and systematicity. There will be quite a number of brute facts, not least of which is the existence of massive quantities of contingent beings: the fundamental particles out of which the physical universe is composed (2015, 60).

And as Richard Swinburne (1979) has argued, theism is not only a far simpler hypothesis than naturalism, God is the simplest kind of person which could possibly exist. God is a person of infinite power, knowledge, and freedom. Limitations in these and any other properties add complexity to the entities that possess them.

Objection 9. *Lack of Proper Causal Relationship Objection*

Michael Murray (2009, 173) formulates this objection as follows:

> Cognitive psychological accounts of religion can account for the origin of religious belief in a way that makes no reference to and requires no causal connection with supernatural reality. However, properly justified belief requires that the target of the belief be causally connected to the belief itself in certain ways. Since these accounts show us that none of those ways is in fact in play in the origins of religious belief, beliefs so generated are unjustified.

This objection can also be formulated in a way that makes it similar to objection number two: natural explanations for religious beliefs show that such beliefs are not properly connected to the object of the belief, and, hence, insensitive to the truth of such beliefs, and, therefore, unjustified. Kim Sterelny (2006), for example, claims that, given the cognitive

mechanisms that gives rise to belief in God, people would believe in God even if there were no object, *i.e.*, God, to which their beliefs were directed. These beliefs would, in orther words, be insensitive to the truth of theism. He formulates the objection as follows:

> Religious commitment cannot both be the result of natural selection for (for example) enhanced social cohesion and be a response to something that is actually divine. A cohesion-and-cooperation model of religion just says that believers would believe, whether or not there was a divine world to which to respond. If a secular theory of the origin of religious belief is true, such belief is not contingent on the existence of traces of the divine in our world. So although a secular and evolutionary model of religion might be (in a strict sense) neutral on the existence of divine agency, it cannot be neutral on the rationality of religious conviction (2006).

Are There Plausible Responses to this Objection?

(A) According to Murray (2009, 175-176), the theist can claim that God is the remote cause of theistic beliefs. He is the best or perhaps the only explanation of a number of features of the world, such as objective morality, why there is something rather than nothing, and the fine-tuning of the universe for life. So, while there are natural explanations of theistic beliefs, there is good independent reason to think that God exists, and, therefore, that He might have implanted in us, so to speak, cognitive mechanisms that give rise to God-beliefs through evolutionary and natural processes.[128]

(B) Clark and Barrett (2011) respond to this objection as follows:

[128] Michael Murray (2009, 174) points out that this would be a problem solely for externalists about epistemic justification, who claim that justifiers are factors outside one's mental life. But, as mentioned in the introductory remarks of this chapter, while evidentialists will say that unreliable faculties can produce justified beliefs – as long as such beliefs are supported by one's evidence, *i.e.*, by one's mental states –, they will also say that once we are apprised of the unreliability of our faculty, its outputs cease to be justified – they are defeated.

> [. . .] God beliefs can be both the result of natural selection and a response to an actual divine person. God beliefs may be justified only if God is the cause of those beliefs. If God is an agent and a person – if God can act and has a will, intention, desires, or goals – then HADD and ToM can put us into the right sort of relation to the object of religious belief. HADD detects agency and ToM detects mind (purpose and intention), so if God is a minded agent, then the god-faculty can produce true beliefs about God (2011, 658-9).

(C) Given the similarities between this objection and objection two, the same response there applies here: insensitivity claims about theistic beliefs face the problem that, if God exists, He is, *pace* Swinburne (2015), most likely a metaphysically necessary being. That is to say that there is no possible world in which He does not exist. The proposition "God exists" is thus a necessary truth. And, since that proposition is true in every possible scenario, insensitivity claims do not work with respect to necessary truths.

Objection 10. *The False Positives Objection*

HADD produces false positives. From this one might conclude that the mechanisms that produce theistic beliefs are error-prone. Such mechanisms are, therefore, unreliable. And if the mechanisms that produce theistic beliefs are not reliable, their outputs cannot be trusted, and, therefore, they cannot be rationally believed (see, for instance, Barrett, 2007; Clark and Barrett, 2011; Murray, 2009).

Are There Plausible Responses to this Objection?

(A) The objection assumes that HADD is the only or at least the predominant cognitive mechanism working in the production of theistic beliefs. But if HADD is only minimally involved in the formation of theistic beliefs, if there are other mechanisms that play a more important role in the formation of such beliefs, as Pascal Boyer (2001) and others have claimed, then the objection has only limited strength.

(B) But, as Clark and Barrett (2011) have argued, even if HADD does play the crucial role in the formation of theistic beliefs that some have claimed, one would need to show that HADD is unreliable in the specific context of formation of belief in God.

(C) And it may also be the case, as Barrett (2007, 68) points out, that there are other mechanisms, such as "our abilities to consider evidence reflectively, [that] can override HADD or any other single cognitive mechanism that tries to generate a belief." If so, the fact that HADD produces false positives would have a limited effect undermining the justification of theistic beliefs.

(D) The objection assumes, as Barrett noted, that "our total system (including HADD) that generalizes belief in intentional agents (such as gods) is error-prone to the extent that it cannot be trusted" (2007, 68). But the problem with this assumption is the difficulty in distinguishing between agencies that have accurately been detected from those that are mere false positives. Purported detection of spirits and ghosts, for instance, can only be identified as false positives if we presuppose that such beings do not exist.

(E) If the human mechanism of agency detection is indeed unreliable, then not only beliefs in gods would be defeated (partially or fully), but all beliefs in intentional agency.

(F) The False Positives Objection is, according to Barrett (2007), self-defeating. The picture of the reliability of our cognitive faculties defended by the False Positives objector can be extended to not only our ability to detect agency, but to our beliefs in general. If the conclusion that our cognitive faculties are error-prone with respect to intentional agency is correct, we can't overlook the more general picture of cognitive limitations that seems to emerge from cognitive science. And if that is the case, it is difficult to avoid systematic collateral damage that will affect the False Positives Objection itself. Thus, the False Positives objector seems, as Barrett put it, to "suffe[r] from suicidal tendencies" (2007, 69). If our cognitive faculties cannot be trusted to give us true beliefs about intentional agents, it most likely cannot be trusted to give us true beliefs about a host of other cognitive domains and will make belief in the False Positives Objection itself unjustified.

(G) Clark and Barrett (2011) argue that it is a mistake to say that HADD and ToM are unreliable mechanisms. They are very effective in the attribution of agency and intention to persons. And even if the objector specifies that they are unreliable in the spiritual domain, all that has been shown, Clark and Barrett (2011, 663) claim, is that they are imprecise or coarse-grained. So, HADD and ToM may be unreliable in producing belief in a specific conception of deity, but not unreliable in producing true belief about some sort of more general supernatural agency.

(H) To say that a cognitive mechanism is unreliable is usually taken to mean that it does not track truth, or that it is insensitive to the truth of the target proposition. In fact, several authors[129] have taken the main objection to the rationality of theistic belief to be understood as a claim to the effect that religious beliefs are insensitive to the truth of the target proposition. So, if theistic beliefs were not true, *i.e.*, if there is no god, then HADD + ToM would still produce religious beliefs. However, as we saw in the responses to the Natural Explanation Objection II and the Lack of Proper Causal Relationship Objection), sensitivity claims in epistemology, and in particular with respect to theism, face several problems. But even if plausible solutions to these problems could be formulated, the fact is, as Joshua Thurow (*e.g.*, 2011) has defended, that religious believers tend to hold their beliefs on the basis of certain reasons and evidence. They will usually claim, for example, that

> they think the Bible is reliable, [that] they think they have witnessed, or know others who claim to have witnessed certain miracles, [that] certain prayers get answered, [that] their life has been changed for the better since believing, [that] the world seems so carefully designed, [that] they've had or know of others who claim to have had religious experience of various kinds, and [that] it is hard to explain all the evidence we have about early Christianity if Jesus wasn't raised from the dead (Thurrow, 2011, 92).

[129] *E.g.*, Clark, and Rabinowitz (2011) and Thurrow (2013).

Thurrow argues that what CSR shows is that religious believers would still have some sort of religious beliefs even if they did not hold their beliefs on the basis of the reasons above. And what the CSR objector would have to show in order to demonstrate that the justification of theistic beliefs has been defeated is that "the processes they actually use, which make use of these kinds of reasons described above, are unreliable" (Thurrow, 2011, 93).

Objection 11. *The Problem of Natural Non-Belief*

Matthew Braddock has advanced what can arguably be claimed to be the most challenging CSR argument against the rationality of theistic beliefs that has appeared in the literature so far. The first step of the argument is the recognition that (a) polytheistic and finite-god beliefs are, from the perspective of theism, false beliefs, that (b) CSR mechanisms have disposed humans to polytheistic and finite-gods beliefs (*i.e.*, polytheists and believers in finite gods are *naturally non-believers* in God), and that (c) throughout history and across a large number of human cultures, humans have been predominantly disposed to such beliefs. In other words, theists have to acknowledge that the mechanisms that produce god beliefs identified by cognitive scientists of religion have produced in the vast majority of cases false beliefs.

The second step tells us that since humans have been predominantly disposed to form false god beliefs, the CSR mechanisms that produce those beliefs are, from a theistic perspective, unreliable. As a result, theists should suspend judgment about the reliability of those mechanisms, unless they have independent evidence suggesting those faculties are reliable. But since theists don't have such evidence, suspension of judgment is the correct response. This, however, should not lead theists to suspend judgment about the existence of God, for there might be other mechanisms or sources of theistic knowledge that are not affected by the unreliability of CSR mechanisms. Consequently, Braddock adds that since CSR mechanisms have contributed significantly to the formation of god beliefs throughout history and across cultures, theists should suspend judgment about the deliverances of those mechanisms, *i.e.*, theists should withhold belief in God, unless they have independent evidence suggesting there are other mechanisms that confer reliability to the beliefs formed by the original CSR mechanism. Given that there is no independent evidence suggesting

there are other mechanisms that confer reliability to the beliefs formed by the original CSR mechanism, the theist has an undercutting reliability defeater for her theistic beliefs formed on the basis of CSR mechanisms that produce God-beliefs, and she can no longer hold her theistic belief rationally (Braddock, 2016, 270-71).

Myron Penner summarizes Braddock's argument as follows (with CSRM standing for "the common belief forming processes that generate god beliefs," and O standing for "factors other than CSRM that influence god beliefs" (2018, 115)):

> (i) Theism must hold that CSRM are massively unreliable;
>
> (ii) Theists have no evidence that O corrects for the unreliability of CSRM;
>
> (iii) One should suspend judgment about outputs of a belief-forming process that is deemed massively unreliable and not corrected for;
>
> (iv) Thus, theists should suspend judgment about the reliability of CSRM and are not justified in accepting some output of CSRM (absent independent evidence for that output) (2018, 115).

Are There Plausible Responses to this Objection?

(A) Both Braddock (unpublished and 2018) and Penner (2018) believe that the fact that CSR mechanism predominantly produce natural non-theistic beliefs doesn't preclude the existence of "approximate" or "vague" reliability in the process that produces theistic beliefs. Penner claims that Braddock "fails to account for two alethically relevant properties – namely, the properties of approximating truths and entailing truths" (2018, 116). The notion of approximating truth stems from the realist/anti-realist debate in philosophy of science. It refers to descriptions of the world (scientific theories) that while not technically true, consist in approximately true descriptions of the world. And Penner believes that the theist can plausibly claim that CSRM deliver approximate truths about the supernatural dimension of reality. Thus, the theist could say that polytheistic beliefs are approximate truths and, consequently, this would deflect the possible epistemically deleterious consequences of the claim that CSR are not fully reliable. And the appeal to the notion of entailing truths (false propositions that entail true ones) can also be helpful for the theist in

that she can claim that false beliefs can still entail truths. For instance: Bob enters the living room and sees Martha's purse on the table. He leaves and five minutes later reenters the living room and no longer sees the purse there. He concludes that Martha was in the room and took her purse. He also infers that he is not alone in the house. But, in fact, it was Silvia, Martha's daughter who took her mother's purse, and, as a result, Bob's belief that Martha took the purse is false. Nonethess, his inferred belief that he is not alone in the house is true. Thus, on Penner's view, the theist can claim that while the CSRM are unreliable in the vast majority of cases, their deliverances can be approximate truths or falsehoods that entail truths.

Braddock argues that CSR research has revealed that we are naturally disposed to believe in the existence of supernatural agents, but not any sort of supernatural agent. Rather, CSR shows that our natural disposition is content biased. As we saw in the introductory chatper of this essay, Barrett (2012b, 322-23) has compiled a list of our core natural religious dispositions or beliefs as identified by CSR. Braddock (2018) uses this list to motivate his response to his own argument (2016) against the rationality of theistic beliefs from CSR. It seems appropriate to recapitulate the thirteen core natural religious intuitions and dispositions identified by CSR before proceeding with our presentation of Braddock's response to his natural non-belief argument against the rationality of theism:

> (A) Elements of the natural world such as rocks, trees, mountains, and animals are purposefully and intentionally designed by someone(s), who must therefore have superhuman power (Kelemen 2004).
>
> (B) Things happen in the world that unseen agents cause. These agents are not human or animal (Guthrie 1993).
>
> (C) Humans have internal components (such as a mind, soul, and/or spirit) that are distinguishable from the body (Bloom 2004, 2007, 2009).
>
> (D) Moral norms are unchangeable – even by gods (Hauser 2006; Katz 2000).
>
> (E) Immoral behavior leads to misfortune; moral behavior to fortune (Jose 1990; Hafer and Begue 2005).

(F) Ritualized behaviors such as marking off special spaces or ritual cleansings can protect from unseen hazards (including those caused by gods) (Liénard and Boyer 2006; Boyer and Liénard 2006).

(G) Some component(s) of humans that has agency (such as souls or minds) may continue to exist without earthly bodies after death (thereby becoming gods) (Cohen and Barrett forthcoming; Bloom 2004).

(H) Gods exist with thoughts, wants, perspectives, and free will to act (Guthrie 1993; Barrett 2012).

(I) Gods may be invisible and immortal, but they are not outside of space and time (Barrett and Keil 1996; Barrett 1999).

(J) Gods can and do interact with the natural world and people, perhaps especially those that are ancestors of the living, and hence, have an interest in the living. This interaction with the world accounts for perceived agency and purpose in the world that cannot be accounted for by human or animal activity (Barrett 2008; Bering 2006, 2002; Boyer 2001).

(K) Gods generally know things that humans do not (they can be super-knowing or superperceiving or both), perhaps particularly things that are important for human relations (Boyer 2001; Barrett and Richert 2003).

(L) Gods, because of their access to relevant information and special powers, may be

responsible for instances of fortune or misfortune; they can reward or punish human actions (Bering and Johnson 2005; Johnson 2005; Boyer 2001; Bering and Parker 2006).

(M) Because of their superhuman power, when gods act, they act permanently, and so when they act in religious rituals, the religious ritual need not be repeated as in baptisms or ordinations (McCauley and Lawson 2002).

These thirteen features of our natural disposition allow Braddock to characterize the content bias of our supernatural disposition as follows:

> humans are disposed to believe in non-human, invisible, disembodied, immortal, super-powerful, super-knowing, super-

> perceiving, infallible, morally interested, punishing/loving, causally active, and minded agents (with beliefs, desires, intentions, character, and free-will) who possess creator or designer status (2018, 8).

Braddock rightly notes (2018, 8) that these dispositions, while not amounting to the maximal attributes of the theistic God, are still theistic-like. We are, in other words, inclined to believe in a theistic-like deity, more so than any other alternative supernatural agent. Hence, CSR research show us that "humans are biased toward vague approximations of [theistic] truths" (2018, 9). This suggests that our cognitive mechanisms responsible for the formation of religious beliefs, while not fully reliable from a theistic perspective, are vaguely or approximately reliable. And contrary to what is claimed in the natural non-belief argument, the vast majority of humans throughout history has not been technically polytheists. For, as Braddock notes (2018, 11, 19-20), the overwhelming majority of humans (more than 99%) have lived their lives since the rise of agricultural societies, 12,000 years ago, with the vast majority living their lives during the period dominated by the Abrahamic religions. The vast majority of humans who have lived have possessed the concept of theistic-like God, with a large number of the total population of this "historic window of theistic progression" (2018, 20) holding either theistic beliefs or close approximations to it. The majority of humans have, from a theistic perspective, held true or approximately true religious beliefs. As a consequence, it would be a mistake to claim that the CSR mechanisms yield mostly false beliefs and that those mechanisms are therefore unreliable.

(B) Another possible response to Braddock's 2016 argument against the rationality of theistic beliefs involves granting Braddock's claim that natural non-belief provides evidence that our god-faculties are (or at least have been in many cases, perhaps in even the majority of them) unreliable, but to claim that this is consistent with at least one variety of theism – Christian theism. One of the central doctrines of Christianity is that God created humanity in His image, with one of the main purposes of having a relationship of trust and love with His creatures. But the creatures rebelled against the Creator, turning their backs on Him and His marvelous offer of an immediate relationship of love with the Being who is the source of all goodness and all beauty. This is called "the Fall." And with the Fall sin (the desire for what is

not good) entered the world. And on many interpretations of the Fall, the human creatures had been endowed with cognitive faculties that enabled them to have direct awareness and understanding of the Creator, but their rebellion and rejection of what is truly good damaged those cognitive faculties, which ceased to function as they were designed to function. In other words, the human noetic system was corrupted by sin and the persistent effects of this corruption or damage is known as the "noetic effects of sin."

So, according to the story of the Fall, God created humans with god-faculties that brought immediate awareness and knowledge of Himself. Human rebellion damaged those faculties and knowledge of God was lost. Hence, on this account of the human predicament, the human God-faculties were originally reliable, but they ceased to be so. God, however, had from the start a plan of salvation and reconciliation of the fallen creatures that would unfold through history, culminating in the incarnation, death, and resurrection of Jesus Christ, the second person of the divine Trinity. This plan of reconciliation involved the restoration of the reliability of the God-faculty damaged by the Fall. This cognitive restoration, however, was (and continues to be) predicated on idea that the human creatures, who chose freely to separate themselves from God, would have to willingly open themselves to a restoration of a relationship of love, trust, and forgiveness with their Creator. As it is the nature of love itself, it cannot be forced on anyone.[130] A loving relationship, by definition, is something that can only emerge from a deliberate act of will of all the parties involved. And, hence, the restoration of the reliability of the human cognitive equipment with respect to knowledge of Himself would have to be accompanied by the restoration of the human will, the will for what is good, primarily for the supreme good which is the development of a loving relationship with He who is the Good Himself. But since this restoration of the human will, of recalibration and reorientation of human desires toward what is truly good for them, is something that cannot be imposed on them (as free creatures, with free *will*, they have to progressively relearn to choose what is good), God can only contribute to this process in an

[130] As Paul Moser explains: "A central divine purpose would be human transformation of a morally significant kind, whereby humans noncoercively become willing to love and to forgive as God loves and forgives, even their enemies, and thereby themselves become personifying evidence of God's reality. The needed volitional transformation would include attunement, or cooperation, of a human will with God's moral will, for the purpose of removing human selfishness and its destructive consequences as a means to building genuine community under divine moral authority" (Moser, 2010, 43).

auxiliary manner, so that that restoration emerges from their own choices, which will cumulatively constitute the process of transformation and restoration of their will and of their God-faculties.[131] As a result, when it comes to this transformation process, God can never impose. He can only invite the creatures to make the right choices. Thus, God's participation in this process has to be subtle and respectful of the person's choices.[132]

And this, I suggest, seems to be a promising direction in which it can be argued that we should expect God to create humans with god-faculties that do not appear to be fully reliable from a theistic perspective, but only vaguely or approximately reliable. To be sure, on this story, our religious faculties were originally fully reliable, became vaguely reliable after the Fall, and their reliability is progressively restored to their full potential as we reenter the sort of relationship God created us to have with Himself, so that our will, desires, and affections are progressively restored, so that we start seeking again what is truly Good. And, again, in a post-Fall world, if God were to simply restore our God-faculties immediately, without a corresponding restoration of our affections, desires, and will, that might turn out to be a small victory, if a victory at all for Him – and for us, for that matter. For if He wants to restore His creatures to a condition of true openness to an everlasting personal relationship of love, trust, and forgiveness, mere propositional knowledge of Himself would be of little help. And in fact, even direct awareness or acquaintance of God, without the proper restoration of our will and soul, would do little to achieve His intended aim for us of giving us

[131] As Dallas Willard put it: "[. . .] in man God had produced a creature that had the responsibility of becoming what he is to become by the choices he makes. God allows, indeed requires, that we choose to act on the basis of our desires, and that we freely decide what we will live for. What we choose in selecting among our desires for fulfillment determines what kinds of persons we become. What we decide to seek in life is the key to our character, and further determines what our character will be. God, like persons in general, wants to be wanted, and tries not to be manifestly present where he is not wanted. He is unwilling to impose himself on anyone if and as long as that can be avoided" (2004).

[132] Here is how C. S. Lewis expressed this line of thought in The Screwtape Letters, by having the senior devil, Screwtape, say the following to his protégé, Wormwood: "You must have wondered why the Enemy does not make more use of his power to be sensibly present to human souls in any degree He chooses and at any moment. But you now see that the Irresistible and the Indisputable are the two weapons which the very nature of His scheme forbids Him to use. Merely to override a human will (as His felt presence in any but the faintest and most mitigated degree would certainly do) would be for Him useless. He cannot ravish. He can only woo. For His ignoble idea is to eat the cake and have it; the creatures are to be one with Him, but yet themselves" (2015, 39).

what is truly the best we could possibly receive, namely, souls truly directed to what is really Good.

A lot more could – and should – be said about this line of reasoning, but I believe that what has been said here suffices to show the general picture of an alternative response to Braddock's 2016 argument, one that pursues a different strategy to account to the (contextualized) unreliability or perhaps vague reliability of our religious faculties given theism.

Objection 12. *The Confirmation Bias Objection*

Some responses to CSR objections, as we have seen, rely on the idea that if the CSR objector wants to show that CSR undermines the rationality of theistic beliefs, even a successful argument for the unreliability of the God-faculty or faculties will not do. The objector will have to show that potential alternative sources of justified theistic beliefs are also unreliable or problematic in some other way. These alternative sources could be religious experiences of different varieties, appearances of design in nature, testimony of reliable religious believers, experiences of miraculous events or testimony of others who claim to have observed such events, arguments from natural theology, answered prayers, the experience of observing the spiritual and moral transformation and growth of committed believers or of oneself, and so on.

The objector could, therefore, attempt to use findings from cognitive psychology to cast doubt on these other potential sources of justified religious belief. De Cruz and de Smedt (2015) and Thurrow (2014a, 2016) have explored the potential undermining effects of certain findings from cognitive psychology to natural theology arguments, and Thurrow (2014b) has explored the potential undermining effects of findings of this sort for religious experiences.

In their book *A Natural History of Natural Theology*, de Cruz and de Smedt (2015) explore the potential defeating effects of certain findings from cognitive psychology to the cosmological argument, the teleological argument, the moral argument, the argument from miracles, and the argument from beauty. They claim that certain versions of moral arguments could potentially be undermined by naturalistic evolutionary accounts of moral

awareness and moral realism. They believe that cognitive psychological research on our intuitions about causality and on our mechanism of agency detection could also potentially pose problems for some versions of the cosmological argument, as they depend on the intuition of the necessity of a first cause and that this cause must be an agent. They also believe empirical research on our tendency to reason teleologically and the way we intuitively assess probabilities "elucidates the lasting popularity of the design argument" (2015, 61), but conclude that certain versions of cumulative design arguments are not undermined by those studies. They found that current research on transmissibility of Minimally Counterintuitive Concepts and on reliance on testimony don't seem to pose any problem for the argument from miracles, and that research on evolutionary accounts of our aesthetic experiences do not undermine the argument from beauty. In summary, de Cruz and de Smedt explore whether cognitive psychological research on the origins of our intuitions could cast doubt on the justifying effects of those intuitions when they are employed in natural theological reasoning, and find that there might be negative repercussions for some versions of some of the traditional arguments from natural theology.

Thurrow (2014a) explores the potential defeating effects of CSR for the cosmological argument, the teleological argument, and C. S. Lewis's argument from desire. He concludes that both the cosmological argument and the teleological argument remain unscathed, but believes that CSR has defeated Lewis' argument from desire.

Lewis states his argument from desire as follows:

> Creatures are not born with desires unless satisfaction for those desires exists. A baby feels hunger: well, there is such a thing as food. A duckling wants to swim: well, there is such a thing as water. Men feel sexual desire: well, there is such a thing as sex. If I find in myself a desire which no experience in this world can satisfy, the most probable explanation is that I was made for another world (1996, 121).

In sum, desires always have corresponding objects that satisfies them. If we have a desire for something transcendent, there must be a transcendent Being with whom we can enter the

sort of relationship that will satisfy our desire for what is transcendent. The problem with Lewis' argument, according to Thurrow, is that CSR theories don't "imply the existence of a desire for something transcendent" (2014a, 286). Thurrow takes this datum to constitute a defeater for Lewis' desire argument.

And what about cosmological and teleological arguments? Can research that shows that we are biased toward teleological and functional explanations defeat the rationality of belief that God exists on the basis of those arguments? Thurrow (2014a, 287-290) explores the possibility of defeat coming from the idea that such beliefs being insensitive to the truth of theism since they would arise, at least in part, from dispositions to reason in terms of purpose. Thurrow believes that this line of reasoning is flawed. He points to the fact that adults with formal education have a reduced propensity to reason teleological as evidence that these dispositions can be counterbalanced. As we undergo formal education, we become better at evaluating evidence and may, according to Thurrow, come to accept those arguments on the basis of the strength of the scientific and philosophical evidence in their favor. Hence, on Thurrow's view, theistic belief on the basis of cosmological and teleological arguments have not been defeated by CSR.

Thurrow (2014b) also explores the potential defeating effects of CSR for religious experiences. He believes that the current state of scientific accounts of religious experiences is too sketchy for any verdict on this question to be issued. Nonetheless, he believes that religious experiences can, *potentially*, be undermined by CSR findings. Perhaps one day scientific research in this area will reach a point in which certain religious experiences can be adequately explained in naturalistic terms, so much so that scientists could accurately predict that once certain natural conditions are realized, such experiences will follow. This, argues Thurrow, could potentially suggest that those experiences are not sensitive to the truth of theism and, thus, theistic beliefs grounded on such experiences could potentially be undermined.

Are There Plausible Responses to this Objection?

(A) While de Cruz and de Smedt believe that certain versions of the moral argument could potentially be undermined by naturalistic evolutionary accounts of moral awareness and

moral realism, they acknowledge that more progress would have to be made in the development of those accounts as they still "seem to be in a sketchy state, whereas theistic accounts (*e.g.*, divine commandment theories) are well developed" (2015, 119). And while they believe that research on our intuitions about causality and on our mechanism of agency detection could potentially pose problems for some versions of the cosmological argument, they believe that, since both our commonsense reasoning and our scientific reasoning rely on these cognitive mechanisms, defeat of the rationality of theistic beliefs grounded on those versions of the cosmological argument could cause collateral damage that would lead to skepticism about the deliverances of our faculties that produce commonsense and scientific beliefs. In other words, appeal to unreliability of our intuitions when it comes to natural theology appears to be self-defeating.

(B) Even if cognitive psychological research on intuitions could put a problem for some versions of the cosmological and teleological arguments, which seems unlikely to occur given the considerations mentioned above, it is unlikely that more recent formulations of those arguments that rely on cosmological data in support of the ideas that the universe had a beginning and that the universe is finely tuned for the existence of intelligent life would be affected be research on intuitions. And even if they would be affected, it is unlikely that scientific reasoning more generally would be spared from the collateral damage of the charge that we can't trust our intuitions with respect to causation, design, teleology, agency detection, and so on. And in such case, self-defeat looms.

(C) The fact that there is disagreement about the effectiveness of natural theological arguments seems to cast doubt on the idea that those moral, causal, teleological, etc., intuitions should be thought to cast doubt on the rationality of beliefs based on those intuitions. De Cruz and de Smedt recognize this potential problem, but they explain it as divergence on the prior probability one has for theism or atheism. According to them, the reason why those intuitions do not convince everyone is because whether they will lead one to find those arguments convincing is because of one's low prior probability of theism. The problem with this solution, however, is that many theists who assign a very high prior probability to theism also find many if not all natural theology arguments unconvincing. So appeal to prior probabilities is unlikely to solve the problem.

(D) The problem with objections to religious experiences from cognitive psychology is fourfold.

(d.1) First, as Thurrow notes, "the science of religious experience is far too young to make any useful judgments on this matter" (2014b, 201).

(d.2) Second, suppose scientific research can eventually adequately explain religious experiences in naturalistic terms with such a precision that scientists will be able to accurately predict the occurrence of certain religious experiences solely on the basis of realization of certain natural conditions. Would that mean that theistic beliefs grounded on such experiences would be defeated. Not necessarily. As Thurrow (2014b, 204) notes, suppose one has independent reasons to believe that God exists (via natural theology, or another set of non-defeated religious experiences, etc.). This would constitute a deflector of the potential defeater constituted by the testimony of the relevant scientists that one's religious experiences of the x kind can be fully explained naturalistically. Thus, the possession of independent evidence to believe theism is true could give the subject reasons to think that religious experiences explained by science are also veridical.

(d.3) Third, it is unclear whether the fact that religious experiences can be explained naturalistically can constitute a defeater for the rationality of holding theistic beliefs on the basis of such experiences, for God could have created the world in such a way that religious experiences are produced via natural means. As Thurrow put it:

> God might well be expected to build into humans some natural way of sensing his presence and feeling as if he is listening. It isn't clear that we should expect God to directly intervene in the physical causal process in these everyday sorts of religious experiences (2014b, 203).

(d.4) Finally, religious experience is a multifaceted and variegated notion. I doubt any science of religious experiences would one day be able to identify (let alone adequately explain in naturalistic terms) all relevant instances of religious experiences. For in the category of religious experiences it should be included not only mystical experiences, but also all sorts of experiences of God's comforting or convicting presence, all sorts of senses

of God's providence in one's life and in the world, all sorts of impressions one can have when reading the Scripture, or when praying, or when watching a sunset, and so on and so forth. Thus, while it is not impossible that scientists may explain naturalistically and predict some kinds of religious experiences, it is highly doubtful that these instances of religious experiences would constitute more than a small fraction of the total number of unique and personal types of religious experiences found in the world.

CONCLUSION

We have come to the end of our survey of the CSR objections to the rationality of theistic beliefs available in the literature. We have seen that there are plausible responses to each of those objections. As a result, the rationality of theistic beliefs doesn't seem to be threatened when it comes to the publicly available evidence with respect to what CSR can say about the rationality of those beliefs. It seems, then, that there are no obvious threats available at the moment to the *ultima facie* rationality of theistic beliefs.

But not only CSR findings don't seem to pose a direct threat to the rationality of theistic beliefs, they seem to be significantly confirmatory of theism in comparison to naturalism. In fact, Braddock argues that theism is much less surprising than naturalism on both the findings of CSR and the existence of natural non-belief. With respect to the first, his core claim (2018, 1) is that

> Pr (supernatural disposition | naturalism) is much lower than Pr (supernatural disposition | theism).

This claim is grounded on the idea that the six features of our supernatural dispositions listed below are more surprising on naturalism than on theism:

> Feature (1) *Theistic Bias:* humans acquire a supernatural disposition that is theistically biased.
>
> Feature (2) *Byproduct:* the right combination of underlying CSR mechanisms evolves by natural selection to collectively generate belief in supernatural agents as a cognitive byproduct.

Feature (3) *Functionless Byproduct:* supernatural belief originates as a byproduct and persists and proliferates throughout the human population despite being costly in terms of fitness.

Feature (4) *Exaptation:* supernatural belief originates as a byproduct but subsequently proliferates because it proves more adaptive than available alternatives.

Feature (5) *Adaptation:* supernatural belief originates and proliferates because it proves more adaptive than available alternatives.

Feature (6) *Theistic Progression:* humans acquire a supernatural disposition that in the course of human history progressively disposes them more to belief in a theistic-like God than to any other alternative supernatural agent(s)—that is, during the monumentally important past 12,000 years when the overwhelming majority of humans (more than 99%) have lived their lives (Braddock, 2018, 12).

Given that our conceptions of supernatural beings can take a large number of different forms, the fact that our natural religious disposition is theistically-biased, as discussed above, is more surprising on naturalism than on theism (feature 1). Why should we, on naturalism, be more disposed to believe in the existence of a being with the omni properties than in a being that lacked those properties? As Braddock notes, "To get a sense of the possibility space, we need only look at the rather large collection of variegated gods that populate the pantheons of known religions" (2018, 13), and to the host of other supernatural beings that have populated the human imaginary (ancestor spirits, angels, demons, ghosts, etc.). On theism, however, we should expect that if God would confer us with dispositions to believe in the existence of supernatural agents, that those dispositions would be to some extent theistically-biased.

On the standard byproduct view of the origins of religious belief (a view that seems to predominate among cognitive scientists of religion), supernatural beliefs did not confer evolutionary advantages, but, rather, as a byproduct of different functional factors that emerged due to random genetic mutation and conferred evolutionary advantage (agency detection, teleological and design reasoning, minimally counterintuitiveness, and so on).

That our disposition to hold supernatural beliefs would emerge in this way also appears to be more surprising on naturalism than on theism (feature 2). As Braddock observes,

> given the plethora of comparably efficient and evolvable variants and how such mechanisms could have easily been modified and combined in various ways during the unguided and fairly contingent course of cognitive evolution, it is surprising given naturalism that humans wound up with the right sort of cognitive architecture— that is, the sort of architecture that would incidentally dispose us to believe in supernatural agents with theistic-like attributes. Given theism, however, this outcome is less surprising. That is, it is not especially surprising that God would guide human cognitive evolution in this direction (2018, 15).

On this predominant account of the origins of supernatural beliefs, religion is not only a byproduct of other functional factors, it is also costly (in the sense that religious behavior requires investment of time and resources in rituals, restrictions of food consumption and sexual practices, etc.). And it is surprising, given naturalism, that supernatural beliefs would persist in our species despite the costs of religious behavior, for it seems that the underlying CSR mechanisms could have been modified in the course of the evolutionary process in order to diminish these costs. Thus, there are evolutionary pressures against the persistence of religious behavior and that religion has persisted despite these constraints is more surprising on naturalism than on theism (feature 3).

The main alternatives to the byproduct theory are the adaptationism model, which claims that religion provided direct evolutionary advantage, and the exaptationism mode, which says that a feature originates as a byproduct but is later co-opted or exapted to serve other functions. On both models, however, the features of our supernatural disposition are highly contingent. Other features could have proved to be more adaptive. For example, on the adaptationist group-selection theory, supernatural beliefs evolved due to the cohesion benefits of their promotion of altruistic behavior. But this role played by supernatural beliefs and behavior as providers of social cohesion could have been played effectively by a large number of different non-theistic contents of supernatural dispositions. Therefore, that

adaptationism or exaptationism is true and we would have the sort of disposition toward supernatural beliefs that are theistically-biased that we have seems to be more surprising on naturalism than on theism (features 4 and 5).

Finally, as seen above, while natural non-belief may seem at first sight to pose a problem for theism in that it would indicate that CSR mechanisms are not reliable, the fact is that the overwhelming majority of humans have been acquainted with theistic-like God concepts. More than 99% of humanity has lived their lives in the past 12,000 years, with the vast majority of humans who have lived their lives during this period having been alive during the prevalence of theistic religions. And this seems to be more surprising on naturalism than on theism (feature 6).

In addition to this cumulative evidential argument in favor of theism, when theism is compared to naturalism with respect to CSR findings, there have been direct challenges to the rationality of atheistic beliefs. Justin Barrett and Ian Church (2013), for instance, have argued that the conjunction of CSR and atheism provides the atheist with a defeater for the reliability of her cognitive faculties in general and that, as a result, one cannot rationally believe both the findings from CSR and atheism. This is so because the cognitive mechanisms that produce religious belief are ordinary cognitive mechanisms that produce beliefs about human minds, the social world, the causal properties of the natural world, and so on.

They formulate the argument as follows:

> 1. There are no gods, souls, and afterlife.
>
> 2. CSR- Belief-Forming Faculties (BFFs)[133] typically produce beliefs in gods, souls, and the afterlife.
>
> 3. Hence, CSR-BFFs typically produce false beliefs.
>
> 4. BFFs that typically produce false beliefs are unreliable.
>
> 5. Hence, CSR-BFFs are unreliable.

[133] CSR-BFFs represent the natural, ordinary belief-forming faculties identified by CSR.

6. Beliefs formed by unreliable faculties lack warrant.

7. CSR-BFFs produce beliefs about human minds (including conscious beliefs, desires, emotions, and their relationship to action), the causal properties of the natural world, and so on.

8. These beliefs (from (7 and 5)) lack warrant (Barrett and Church, 2013, 7 of authors' draft).

In other words, unreliability claims about the cognitive mechanisms that produce religious beliefs would have negative epistemic implications for a vast range of non-religious beliefs, including beliefs about CSR and its findings and about atheism. In claiming that the BFFs that produce religious beliefs are unreliable, the atheist would be committed to skepticism about the reliability of the cognitive faculties that produce belief in CSR and atheism itself. As Barrett and Church put it, "atheism, plus the accounts from CSR provide reason to doubt the belief-forming faculties relevant for forming beliefs about atheism and about CSR" (Barrett and Church, 2013, 9).

Barrett and Church identify one possible objection to their argument as coming from a criticism of premise 3. The atheist could attempt to claim that the unreliability of the faculties when forming religious beliefs cannot be generalized to non-religious contexts. The problem with this response, however, is that the atheist would have to show how the belief formation in these two contexts differs in their reliability, a task that doesn't seem feasible since belief formation in these two contexts seems indistinguishable. And Myron Penner (2018) criticizes Barrett and Church's argument on the basis of their claim that in order to escape self-defeat the atheist who believes in atheism on the basis of CSR findings would have to show that BFFs used for God beliefs and beliefs about other agents are relevantly dissimilar. Penner believes the atheist can meet this challenge by saying that "human agents are accessible in ways that gods, if any there be, are not" (127). In other words, the atheist could, according to Penner, claim that differences in our epistemic distance to God and to other agents make the BFFs involved in beliefs about divine agency and human agency relevantly dissimilar.

Another challenge for atheism coming from the work of cognitive psychologists originates from studies that show very significant statistical correlation between autism and atheism. But more than that, atheism seems to be mediated by autism (Clark, 2014). Autistic people lack ToM, that is to say, they have "mentalizing deficits," not immediately recognizing people's beliefs, feelings, and desires. And religious belief is, as we have seen, produced by a combination of cognitive mechanisms that produce agency detection and attribution of intentions and purposes to such agents. So religious belief would involve properly functioning cognitive faculties and atheism is correlated with and mediated by a cognitive defect. Therefore, belief in atheism is not formed in an epistemically justified or rational way.

The authors of one of the studies that show a strong correlation between autism and atheism suggest that these findings should be treated with caution: "We emphasize that our data do not suggest that disbelief solely arises through mentalizing deficits; multiple psychological and socio-cultural pathways likely lead to a complex and overdetermined phenomenon such as disbelief in God" (2012, 5). And Kelly Clark claims that the findings from those studies merely show that atheism is abnormal (2013). For Clark, "Given the complexities of both the human mind and human culture, it is impossible to tell [whether atheism is irrational]" (2014).

In conclusion, in this chapter we have seen that the justification of theistic beliefs seem to be safe from genuine and ultimate defeat when it comes to the findings of CSR. But more than escape safely from defeat, theism seems to emerge from the current debate on the epistemological implications of CSR in a more favorable position than naturalism/atheism. On the one hand, the six features of our supernatural dispositions that we have just seen seem to provide a powerful cumulative evidential case in favor of theism in comparison with naturalism. On the other hand, there are challenging objections to the justification of atheistic beliefs coming from CSR and cognitive psychology more generally, namely, the correlation between atheism and autism and the charge that atheists cannot use the findings from CSR to undermine the rationality of religious beliefs without incurring in self-defeat. Surely, this is just the beginning of a fascinating debate that promises to make the

dialogue between cognitive scientists and philosophers interested in the religious questions a most vibrant and enlightening one for years to come.

CONCLUSION

We have come to the end of our exploration of the epistemological implications of certain findings from cognitive science. It's time then to recapitulate our results.

We set out to examine two claims: first, that cognitive science indicates that evidentialism is false, and, relatedly, that cognitive science of religion (CSR) indicates that religious evidentialism is false; and, secondly, that CSR findings pose a problem for the rationality of theistic beliefs. The first part of this essay focused on the first claim, and the second part explored the second claim. In order to properly evaluate the first claim, we began with an exploration of the current literature on the nature of the evidence and on evidentialism. We then explored some of the main findings from cognitive science relevant for our evaluation of the first claim. With the results of these three chapter in place, we moved to the evaluation of the first claim. With respect to the second claim, we began by laying out the traditional understanding of rationality as closely related to the notion of justification, and briefly discussed the important notion of total evidence. We then explored the nature of defeaters and identified a wide variety of types of defeaters that have been employed in epistemological debates, as well as some of the main challenges to the notion of defeat and how certain evidentialist understandings of it seem to be immune from those criticisms. With this general picture of what rationality and defeaters are and of some of the main ways in which they have been employed in epistemological debates, we examined twelve objections to the rationality of theism.

So, what our exploration of these questions can tell us about the compatibility of evidentialism with cognitive science and whether or not there are defeaters for the rationality of theistic beliefs from CSR? Let us recapitulate in more detail what we did in each chapter, so that we can see more clearly how our exploration of evidence, evidentialism, CSR findings, rationality, and defeaters bear on the two claims.

First, we presented the contemporary debate on the nature of evidence and on the viability of an evidentialist understanding of epistemic justification. We explored five major understandings of what evidence, ontologically speaking, is, and the four main roles the notion of evidence is expected to play. We saw that, of these understandings of the ontology

and roles of evidence, evidentialists have favored the view that evidence consists of mental states (experiences, beliefs, dispositions, and so on), and that the primary roles evidence is expected to play are to justify beliefs and to be what rational subjects seek in forming beliefs. And we saw that, despite the many challenges that have been put forward against their views, evidentialists have been able to formulate a coherent and explanatorily rich understanding of the evidentialist thesis, one that can satisfy the main considerations of the evidentialist schema and offer plausible responses to the main challenges that have been presented against evidentialism. We then discussed a number of experiments conducted by cognitive and developmental psychologists whose results have given rise to the naturalness thesis. After exploring different understandings of this thesis, we presented three major understandings of the cognitive mechanisms involved in the origins of religious beliefs (the attributional, dispositional, and preparedness approaches). The final chapter of the first part sought to answer the first claim, in part, by evaluating whether these three main understandings of the origins of religious beliefs were compatible with an evidentialist/mentalist understanding of epistemic justification. It also sought to answer the first claim by evaluating three arguments from cognitive science (by Greco, McCauley, and Barrett) directly targeting the evidentialist thesis. We saw that not only does evidentialism have the resources to escape those objections and to deliver the result that religious beliefs formed on the basis of the three main CSR mechanisms can be justified, it also has the resources to provide adequate models of the *sensus divinitatis* and of reformed epistemology – and which are arguably superior to the non-evidentialist models available.

In the second part, we began by exploring the traditional understanding of rationality and saw that there is some plausibility to a non-traditional understanding that separates rationality from justification and reasonableness, but we acknowledged that more work needs to be done on this front before a verdict can be issued on which conception of rationality-justification-reasonableness debate is the correct one. We then explored what defeaters are, what types of defeaters there are, and how they can be operationalized. We focused in particular on Alvin Plantinga's understanding and typology of defeaters, since the circumstances and dialectic that gave rise to his formulation of his account of defeaters – his response to his criticisms of his Evolutionary Argument against Naturalism – are similar to

those found in many CSR objections to the rationality of theistic beliefs. We then presented two objections to the traditional notion of epistemic defeat and saw that some versions of evidentialism explored in the chapters of the first part seem to have the resources to escape those objections. And, finally, we presented twelve objections to the rationality of theistic beliefs from CSR and saw that there are plausible responses to each of those objections. So, while these objections may in some cases (misleadingly) defeat the justification for theistic beliefs, the evidence we have at the moment suggests that they are unlikely to provide genuine defeat for the justification one has for theistic beliefs. But more than that, more than escape safely from defeat, theism seems to emerge from the current debate on the epistemological implications of CSR in a more favorable position than naturalism/atheism. For, as we saw in the final chapter, CSR seems to provide a powerful cumulative evidential case in favor of theism in comparison with naturalism, and there are challenging CSR objections against the rationality of atheism.

In conclusion, cognitive science and, in particular, CSR are young disciplines. Naturally, much more work – both empirical and theoretical – lies ahead of scientists working in these areas before much of the central debates to which they have given rise, including those of philosophical significance, can be settled. Their findings are, for the most part, amenable to falsification and revision. Still, much progress has been made on a number of fronts by these disciplines. Philosophers, in particular epistemologists, do well to pay close attention to these debates and try to draw possible philosophical implications from them. This is what we attempted to do in this essay. It is our hope that it can contribute to further debate and further interaction between cognitive scientists and philosophers that can be mutually beneficial.

REFERENCES

Alston, W. (1988). An internalist externalism. *Synthese*, 74, 265-283.

Audi, R. (1994). Dispositional beliefs and dispositions to believe. *Nous*, 28.4, 416–434.

Audi, R. (2004). Theoretical rationality: Its source, structure, and scope. In Mele, A. & Rawling, P. (eds.) *The oxford handbook of rationality*. Oxford, UK: Oxford University Press.

Audi, R. (2011). *Rationality and religious commitment*. Oxford, UK: Oxford University Press.

Audi, R. (2014). Intuition, agency detection, and social coordination as analytical and explanatory constructs in the cognitive science of religion. Trigg, R. &. Barrett, J. (eds.) *The roots of religion: exploring the cognitive science of religion*. Farnham, UK: Ashgate.

Arnold, A. (2013). Some evidence is false. *Australasian Journal of Philosophy*, 91 (1), 165 – 172.

Atran, S. (2002). *In Gods we trust: The evolutionary landscape of religion*. Oxford: Oxford University Press.

Banerjee, K., & Bloom, P. (2013). Would Tarzan believe in God? Conditions for the emergence of religious belief. *Trends in Cognitive Sciences*, January, Vol. 17, No. 1

Banerjee, K., & Bloom, P. (2014). Why did this happen to me? Religious believers' and non-believers' teleological reasoning about life events. *Cognition,* 133, 277–303.

Banerjee, K., & Bloom, P. (2015). "Everything happens for a reason": Children's beliefs about purpose in life events. *Child Development*, March/April, Volume 86, Number 2, 503–518.

Banerjee, K. & Bloom, P. (2017). You get what you give: Children's karmic bargaining. *Developmental Science*. Volume 20, Issue 5, September.

Baker-Hitch, M., & Benton, M. (2015). Defeatism defeated, *Philosophical Perspectives*: Volume 29, Issue 1, 40-66.

Barrett, J. (2004). *Why would anyone believe in God?* Berkeley, CA: Alta Mira Press.

Barrett, J. (2007). Is the spell really broken: Biopsychological explanations of religion and theistic belief. *Theology and Science*, Volume 5, Issue 1, 61-2.

Barrett, J. (2008). Why Santa Claus is not a God. *Journal of Cognition and Culture*, 8, 149–161.

Barrett, J. (2009). Cognitive science, religion, and theology, In Schloss, J., Murray, M. (eds.) *The believing primate: Scientific, philosophical, and theological reflections on the origin of religion*. Oxford, UK: Oxford University Press.

Barrett, J. (2011). *Cognitive Science, Religion, and Theology: From Human Minds to Divine Minds*. Philadelphia, PA: Templeton Press.

Barrett, J. (2012a) *Born Believers: The Science of Children's Religious Belief.* Free Press.

Barrett, J. (2012b). Towards a cognitive science of Christianity. In Stump, J & Padgett, A. (eds.), *The Blackwell Companion to Science and Christianity*. Oxford: Blackwell Publishing.

Barrett, J., Richert, R., & Driesenga, A. (2001). God's beliefs versus mother's: The development of nonhuman agent concepts. Volume 72, Issue 1, 50-65.

Barrett, J. & Lanman, J. (2008). The science of religious beliefs. *Religion,* 38, 109-124.

Barrett, J. & Church, I. (2013). Should CSR give atheists epistemic assurance? On beer-goggles, BFFs, and skepticism regarding religious beliefs, *The Monist*, Vol. 96, No. 3, July, 311-324.

Bergmann, M. (1997). Internalism, externalism, and the no-defeater condition. *Synthese*, 110: 399-417.

Bergmann, M. (2007). *Justification without awareness: A defense of epistemic externalism*, Oxford, UK: Oxford University Press.

Bergmann, M. (2008). Reidian externalism. In Hendricks, V. (ed.) *New waves in epistemology*. Basingstoke, UK: Palgrave-Macmillan.

Bergmann, M. (2009). Rational disagreement after full disclosure. *Episteme*. Volume 6, Issue 3, 336-353.

Bergmann, Michael (2013). Phenomenal conservatism and the dilemma for internalism. In Tucker, Chris (ed.), *Seemings and justification: New essays on dogmatism and phenomenal conservatism.* Oxford, UK: Oxford University Press.

Bergmann, M. (2014). Rational religious belief without arguments. In *Philosophy of Religion: An Anthology*, Rea and Pojman (eds.) (7th ed.) Boston, MA: Wadsworth.

Bering, J. (2002). The existential theory of mind. *Review of General Psychology*, Vol. 6, No. 1, 3–24.

Bering, J. (2006). The folk psychology of souls. *Behavioral and Brain Sciences*, 29, 453–498.

Bering, J. (2011). *The belief instinct: The psychology of souls, destiny, and the meaning of life*. New York, NY: Norton.

Bering, J., Blasi, C., & Bjorklund, D. (2005). The development of afterlife beliefs in religiously and secularly schooled children. *British Journal of Developmental Psychology*, 23(4), 587-607.

Bering, J. & Bjorklund, D. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology*, 40(2), 217-33.

Bloom, P. (2004). *Descartes' baby: How the science of child development explains what makes us human*. New York, NY: Basic Books.

Bloom, P. (2007). Religion is natural. *Developmental Science*, 10:1, 147–151.

Bonjour, Lawrence (1998) *In Defense of Pure Reason*. Cambridge University Press.

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York, NY: Basic Books.

Bonjour, L. (1998). *In Defense of Pure Reason*. Cambridge, MA: Cambridge University Press.

Bonjour, L. (2003). *A version of internalist foundationalism*. In BonJour, L. & Sosa, E., Epistemic justification: Internalism vs. externalism, foundations vs. virtues, Malden, MA: Blackwell.

Braddock, M. (2016). Debunking arguments and the cognitive science of religion. *Theology and science*, 14 (3), 268-287.

Braddock, M. (2018). An evidential argument for theism from the cognitive science of religion. In Eyghen, H., Peels, R., van den Brink, G. (eds.) *New developments in the cognitive science of religion: The rationality of religious belief*. New York, NY: Springer International Publishing.

Braddock, M. (unpublished) Theism, naturalism, and the problem of natural non-belief.

Brewer, W., Chinn, C., & Samarapungavan, A. (2000). Explanation in scientists and children. In Keil, F. & Wilson, R. (eds.) *Explanation and cognition*. Cambridge, MA: MIT Press.

Byerly, T. R. (2012). The evidential support relation in epistemology. *Oxford bibliographies online – philosophy section*. Pritchard, D. (ed.). Retrieved from http://www.oxfordbibliographies.com/view/document/obo-9780195396577/obo-9780195396577-0189.xml

Byerly, T. R. (2013). Explanationism and jutified beliefs about the future. *Erkenntnis* 78, 229-243.

Byerly, T. R. (2014). A dispositional internalist evidentialist virtue epistemology. *Logos and Episteme* 5, 4, 399-424.

Byerly, T. R. & Martin, K. (2014). Problems for explanationism on both sides. *Erkenntnis* 80 (4), 773-791.

Byerly, T. R. & Martin, K. (2016). Explanationism, super-explanationism, ecclectic explanationism: Persistent problems on both sides. *Logos and Episteme,* 7 (2), 201-213.

Callan, M., Ellard, J., Nicol, J. (2006). The belief in a just world and immanent justice reasoning in adults. *PSPB*, Vol. 32 No. 12, 1646-1658.

Carter, J. (2015). Group knowledge and epistemic defeat. *Ergo*. 2(28), 711-735

Choe, K., Keil, F., & Bloom, P. (2011). Developing Intuitions about How Personal and Social Properties Are Linked to the Brain and the Body. *Infant and Child Development,* 221: 430-441.

Clark, K. & Rabinowitz, D. (2011). Knowledge and the objection to religious belief from cognitive science. *European Journal for Philosophy of Religion*, 3, 67-81.

Clark, K. & Barrett, J. (2010). Reformed epistemology and the cognitive science of religion. *Faith and Philosophy*, Volume 27, Issue 2, 174-189.

Clark, K. & Barrett, J. (2011). Reidian religious epistemology and the cognitive science of religion. *Journal of the American Academy of Religion*, vol. 79, n. 3.

Clark, K. (2013). *Is atheism 'normal'? Reflections from the cognitive science of religion*. Retrieved from https://www.youtube.com/watch?v=xgT8LHyKIak

Clark, K. (2014). Is atheism irrational? *Big Questions*. Retrieved from https://www.bigquestionsonline.com/2014/01/28/atheism-irrational/

Climenhaga, N.(2017). Intuitions are used as evidence in philosophy. *Mind*. Online First Version.

Comesaña, J. (2010). Evidentialist reliabilism. *Noûs* 44.4, 571–600.

Comesaña, J. & Kantin, H. (2010). Is Evidence knowledge? *Philosophy and Phenomenological Research*. Volume 80, Issue 2, 447–454.

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

Conee, E. & Feldman, R. (2004). *Evidentialism: Essays in epistemology*, New York: Oxford University Press.

Conee, E. & Feldman, R. (2008). Evidence. In Smith, Q. (ed.), *Epistemology: New Essays.* Oxford, UK: Oxford University Press.

Conee, E., & Feldman, R. (2011). Replies. In Dougherty, T. (ed.), *Evidentialism and its discontents*. Oxford, UK: Oxford University Press.

Chisholm, R. (1977). *Theory of knowledge* (second edition). Englewood Cliffs, NJ: Prentice Hall.

Dancy, J. (2000). *Practical reality*. Oxford, UK: Oxford University Press.

Dawkins, R. (2008). *The God delusion*. Boston, MA: Houghton Mifflin Harcourt.

de Almeida, C. & Fett, J. (2015). Defeasibility and gettierization: A reminder. *Australasian Journal of Philosophy*, 94, 152–69.

de Almeida, C. (2018). Knowledge, benign falsehoods, and the gettier problem. In *Explaining knowledge: New essays on the Gettier problem*. Borges, R., de Almeida, C., Klein, P. (eds.). Oxford, UK: Oxford University Press.

de Cruz, H. & de Smedt, J. (2015). *A natural history of natural theology*. Cambridge, MA: MIT Press.

Dennett, D. (2007). *Breaking the spell: Religion as a natural phenomenon*. London, UK: Penguin Books.

DePoe, J. (n.d.). Knowledge by acquaintance and knowledge by description. *International Encyclopedia of Philosophy*. Retrieved from http://www.iep.utm.edu/knowacq/

DePoe, J. (2011). Defeating the self-defeat argument for phenomenal conservativism. *Philosophical Studies*, 152 (3):347-359.

DePoe, J. (2012). Bergmann's dilemma and internalism's escape. *Acta Analytica* 27 (4), 409-423.

Dougherty, T. (2011). In Defense of propositionalism about evidence. In Dougherty, T. (ed.), *Evidentialism and its discontents*. Oxford, UK: Oxford University Press.

Dougherty, T. (ed.) (2011b). *Evidentialism and its discontents*. Oxford, UK: Oxford University Press.

Dougherty, T. & Rysiew, P. (2013). Experience first. In Steup, M. & Turri, J. (eds.), *Contemporary debates in epistemology*. Oxford, UK: Blackwell.

Dougherty, T. and Tweedt, C. (2015). Religious epistemology. *Philosophy Compass.* Volume10, Issue 8, 547-559.

Dougherty, T. & Gage, L. (2015). New atheist approaches to religion. In Oppy. G. (ed.), *The routledge handbook of contemporary philosophy of religion*. New York: Routledge. pp. 51-62.

Evans, M. (2001) Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology*, 42, 217-266.

Fales, E. (1996) *A defense of the given*. Lanham, MD: Rowman & Littlefield.

Fantl, J., Mcgrath, M. (2009). *Knowledge in an uncertain world*. Oxford, UK: Oxford University Press.

Feldman, R., and Conee, E. (2001). Internalism defended. *American Philosophical Quarterly*, Vol. 38, No. 1, pp. 1-18.

Feldman, R. (2003). *Epistemology*. Englewood Cliffs, NJ: Prentice Hall.

Feldman, R. (2004). Having evidence. In Conee, E. & Feldman, R. *Evidentialism*. Oxford, UK: Oxford University Press.

Frise, M. (2015). Memory, epistemology of. *Internet Encyclopedia of Philosophy*. Retrieved from http://www.iep.utm.edu/epis-mem/

Frise, M. (2017a). Preservationism in the epistemology of memory. *Philosophical Quarterly*, 67 (268).

Frise, M. (2017b). Internalism and the problem of stored beliefs. *Erkenntnis*, 82 (2), 285-304.

Frise, M. (2018). Eliminating the problem of stored beliefs. *American Philosophical Quarterly*, 55 (1), 63-79.

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

Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23(6): 121–123.

Gelman, A., and Kremer, E. (1991). Understanding natural cause: Children's explanations of how objects and their properties originate. *Child Development*, 62(2), 396-414.

Gergely, G., Nádasdy, Z., Csibra, G. , & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56, 165-193.

Ginet, C. (1995). Comments on Plantinga's two-volume work on warrant. *Philosophy and Phenomenological Research*, Vol. LV, no. 2.

Goldberg, S. (2015). The asymmetry thesis and the doctrine of normative defeat. *Social and epistemic norms conference, the philosophy and theology of intellectual humility project*, Saint Louis University. Retrieved from https://www.youtube.com/watch?v=P2ZJNZIVlmY.

Goldman, A. (1999). Internalism exposed. *The Journal of Philosophy*, 96, 271-93.

Goldman, A. (2009). Williamson on knowledge and evidence. In Greenough, P., Pritchard, D., & Williamson, T. (eds.), *Williamson on knowledge*. Oxford, UK: Oxford University Press.

Goldman, A. (2011). Toward a synthesis of reliabilism and evidentialism? Or: Evidentialism's troubles, reliabilism's rescue package. In Dougherty, T. (ed.) *Evidentialism and its discontents*. Oxford, UK: Oxford University Press.

Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system. In Keil, F. & Wilson, R. (eds.), *Explanation and cognition*. Cambridge, MA: MIT Press.

Gottfried, G., Gelman, S., & Schultz, J. (1999). Children's understanding of the brain: from early essentialism to biological theory. *Cognitive Development*, 14(1), 147-174.

Greco, J. (2006). *Achieving knowledge: A virtue-theoretic account of epistemic normativity*. Cambridge, UK: Cambridge University Press.

Greco, J. (2011). Evidentialism about knowledge. In Dougherty, T. (ed.) *Evidentialism and its discontents*. Oxford, UK: Oxford University Press.

Guthrie, S. (1980). A cognitive theory of religion. *Current Anthropology*, Vol. 21, No. 2, Apr., pp. 181-203.

Hasan, A. (2013). Phenomenal conservatism, classical foundationalism, and internalist justification. *Philosophical Studies*, 162, 119–141.

Hasan, A. & Fumerton, R. (2014). Knowledge by acquaintance vs. description. *Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/knowledge-acquaindescrip/

Harman, G. (1973). *Thought*. Princeton, NJ: Princeton University Press.

Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford, UK: Oxford University Press.

Hornbeck, R., Barrett, J., Kang, M. (eds.) (2017) Religious cognition in China "homo religiosus" and the dragon. New York, NY: Springer International Publishing.

Huemer, M. (1999). The problem of memory knowledge. *Pacific Philosophical Quarterly*, 80, 346–57.

Huemer, M. (2001). *Skepticism and the veil of perception*. Rowman & Littlefield.

Huemer, M. (2005). *Ethical intuitionism*. Basingstoke, UK: Palgrave-Macmillan.

Huemer, M. (2007). Compassionate phenomenal conservatism. *Philosophy and Phenomenological Research,* 74 (1), 30–55.

Huemer, M. (2013a). Phenomenal conservative. *Internet Encyclopedia of Philosophy*. Retrieved from http://www.iep.utm.edu/phen-con/

Huemer, M. (2013b). Phenomenal conservatism Über Alles. In Tucker, Chris (ed.) *Seemings and justification: New essays on dogmatism*. Oxford, UK: Oxford University Press.

Hume, D. (1748). *An enquiry concerning human understanding*. Gutenberg Project. Retrieved from https://socialsciences.mcmaster.ca/econ/ugcm/3ll3/hume/enquiry.pdf

Järnefelta, E., Zhub, L., Canfieldc, C., & Kelemen, D. (2015). The divided mind of a disbeliever: Intuitive beliefs about nature as purposefully created among different groups of non-religious adults. *Cognition*, 140, 72–88.

Järnefelta, E., Zhub, L., Canfieldc, C., Chena, M., & Kelemen, D. (2018). Reasoning about nature's agency and design in the cultural context of China. *Religion, Brain and Behavior*. Online first article.

Johnson, C. (1990). If you had my brain, where would I be? Children's understanding of the brain and identity. *Child Development*, Vol. 61, No. 4, pp. 962-972.

Johnson, D. (2011). Proper function and defeating experiences. *Synthese*, volume 182, issue 3.

Keil, F. (1989). Concepts, kinds, and cognitive development. Cambridge, MA: MIT Press.

Keil, F. and Wilson, R. (2000). Explaining explanation. In Keil, F. & Wilson, W. (eds.), *Explanation and cognition*. Cambridge, MA: MIT Press.

Kelemen, D. (1999a). Why are rocks pointy? Children's preference for teleological explanations of the natural world. *Developmental Psychology*, 35, 1440–1452.

Kelemen, D. (1999b). The scope of teleological thinking in preschool children. *Cognition,* 70, 241–272.

Kelemen, D. (2004). Are children intuitive theists? Reasoning about purpose and design in nature. *Psychological Science*, 15, 295–301.

Kelemen, D., & Rosset, E. (2009). The human function compunction: Teleological explanation in adults. *Cognition*, 111, 138 –143.

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default, Journal of Experimental Psychology: General, Vol. 142, No. 4, 1074–1083.

Kelly, T. (2006). Evidence. *Stanford Encyclopedia of Philosophy.* Retrieved from https://plato.stanford.edu/entries/evidence/

Kelly, T. (2008). Evidence: Fundamental concepts and the phenomenal conception. *Philosophy Compass*. 3 (5), 933-955.

Klein, P. (1981). *Certainty: A refutation of skepticism*. Minneapolis, MN: Minnesota University Press.

Klein, P. (2010). Self-Profile, In Dancy, J., Sosa, E., & Steup, M. (eds.), *A Companion to epistemology* (Second Edition). Oxford, UK: Wiley-Blackwell.

Kotzen, M. (2010). A formal account of epistemic defeat. (unpublished manuscript). Retrieved from http://fitelson.org/seminar/kotzen.pdf

Kvanvig, J. (1989). Conservatism and its virtues, *Synthese*, 79, 143-63.

Kvanvig, J. (ed.) (1996) *Warrant in contemporary epistemology: Essays in honor of Plantinga's theory of knowledge*. Latham, MD: Rowman & Littlefield Publishers.

Kvanvig, J. (2005). A Perspective on Plantinga's theory of warrant. *Certain Doubts*, Retrieved from http://certaindoubts.com/a-perspective-on-plantingas-theory-of-warrant/

Kvanvig, J. (2007). Two approaches to epistemic defeat, In *Alvin Plantinga: Contemporary philosophy in focus*, Baker, D. (ed.), Cambridge, UK: Cambridge University Press.

Kuhlmeier, V., Bloom, P. & Wynn, K. (2004). Do 5-month-old infants see humans as material objects? *Cognition,* 94, 95–103.

Lackey, J. (2008). *Learning from words: Testimony as a source of knowledge*, Oxford, UK: Oxford University Press.

Lasonen-Aarnio, M. (2010). Unreasonable knowledge. *Philosophical Perspectives* 24: 1–21.

Leech, D. & Visala, A. (2011a). The cognitive science of religion: A modified theist response. *Religious Studies*, 47, 301–316.

Leech, D. & Visala, A. (2011b). The cognitive science of religion: Implications for theism? *Zygon*, 46, 47–65.

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

Lewis, C. S. (2007). Mere Christianity. In The complete C. S. Lewis signature classics. New York, NY: Harper Collins.

Lewis, C. S. (2015). *The Screwtape letters*. New York, NY: Harper Collins.

Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, 74: 549-67.

Lillard, A. (1996). Body or mind: Children's categorizing of pretense. Volume 67, Issue 4: 1717-1734.

Littlejohn, C. (2011). Defeating phenomenal conservatism. *Analytic Philosophy.* 52, 35-48.

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

Littlejohn, C. (2013). No evidence Is false. *Acta Analytica*, 28 (2), 145-159.

Locke, J. (1690). *An essay concerning human understanding*. Gutenberg Project. Retrieved from http://central.gutenberg.org/eBooks/WPLBN0000660463-An-Essay-Concerning-Human-Understanding-by-Locke-John.aspx?&Words=An%20essay%20concerning%20human%20understanding

Long, T. (2015). Review of Kevin McCain's book (2014) Evidentialism and epistemic justification, Routledge. In *Notre Dame Philosophical Reviews*. Retrieved from https://ndpr.nd.edu/news/evidentialism-and-epistemic-justification/

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science*, 14(2), 453-460.

Lycan, W. (1988). *Judgment and justification*. Cambridge, UK: Cambridge University Press.

Lyons, J. (2011). Circularity, reliability, and the cognitive penetrability of perception, *Philosophical Issues.* 21 (1), 289-311.

Lyons, J. (2016). Experiential evidence? *Philosophical Studies*, 173 (4), 1053-1079.

Mawson, T. (2014). Cognitive science of religion and the rationality of classical theism. In Trigg, R. & Barrett, J. (eds.) *The roots of religion*. Farnham, UK: Ashgate.

McAllister, B. (2018). Seemings as sui generis. *Synthese*, Volume 195, Issue 7, 3079–3096.

McAllister, B. & Dougherty, T. (2018) Reforming reformed epistemology: A new take on the sensus divinitatis. *Religious Studies.* Published online: 07 May 2018.

McCain, K. (2008). The virtue of epistemic conservatism. *Synthese*, 164, 185-200.

McCain, K. (2014a). A New evil demon? No problem for moderate internalists. *Acta Analytica*.

McCain, K. (2014b). *Evidentialism and epistemic justification*. New York, NY: Routledge.

McCain, K. (2015). Explanationism: defended on all sides. *Logos and Episteme,* 6 (3), 333-349.

McCain, K. (2017). Undaunted explanationism. *Logos and Episteme*, 8 (1), 117-127.

McCain, K. (2018). Explanationist aid for phenomenal conservatism. *Synthese*, 195 (7), 3035-3050.

McCauley, R. (2000). Naturalness of religion, unnaturalness of science. In Keil, F. & Wilson, R. (eds.), *Explanation and cognition*. Cambridge, MA: MIT Press.

McCauley, R. (2011) *Why religion is natural and science Is not.* Oxford, UK: Oxford University Press.

McGrew, T. (1995). The Foundations of Knowledge, Lanham, MD: Littlefield Adams Books.

McGrew, T. (2011). Evidence. In Bernecker, S. & Pritchard, D. (eds.) *The Routledge Companion to Epistemology*. New York, NY: Routledge.

Meltzoff, A., & Moore, K. (1983). Newborn infants imitate adult facial gestures. *Child Development,* Vol. 54, No. 3, 702-709.

Murray, M. (2009). Scientific explanations and the justification of religious belief, In Schloss, J. & Murray, M. (eds.) *The believing primate: Scientific, philosophical, and theological reflections on the origin of religion*. Oxford, UK: Oxford University Press.

Mitova, V. (2015). Truthy psychologism about evidence. *Philosophical Studies*, 172 (4), 1105-1126.

Moon, A. (2012a). Three forms of internalism and the new evil demon problem. *Episteme*, 9, 345–360.

Moon, A. (2012b). Knowing without Evidence. *Mind,* 121, 309–331.

Moon, A. (2015) The new evil demon, a frankfurt-style counterfactual intervener, and a subject's perspective objection: Reply to McCain. *Acta Analytica* 30 (1), 107-116.

Moon, A. (2016). Recent work in reformed epistemology. *Philosophy Compass*, 11, 879–891.

Moon, A. (2018). Evidentialism, time-slice mentalism, and dreamless asleep. In *Believing in accordance with the evidence: New essays on evidentialism*, McCain, K. (ed.). New York, NY: Springer International Publishing.

Moretti, L. (2015). Phenomenal conservatism. *Analysis,* 75 (2), 296-309.

Morton, J., & Johnson, M. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, Vol. 98, No. 2, 164-181.

Moser, P. (1989). *Knowledge and evidence*. Cambridge, UK: Cambridge University Press.

Moser, P. (2008). *The eusive God: Reorienting religious epistemology*. Cambridge, UK: Cambridge University Press.

Moser, P. (2010). *God and evidence: Religious knowledge reexamined*. Cambridge, UK: Cambridge University Press.

Neta, R. (2008). What evidence do you have? *British Journal for the Philosophy of Science*, 59 (1), 89-119.

Neta, R. (2018). Your evidence is the set of facts that are manifest to you. in *The factive turn in epistemology*. (ed. Velislava Mitova), Cambridge, UK: Cambridge University Press.

Newman, G., Keil, F., Kuhlmeier, V., & Wynn, K. (2010). Early Understandings of the link between agents and order. *PNAS*, 107 (40), 17140-17145.

Norenzayan, A., Lee, A. (2010). It was meant to happen: Explaining cultural variations in fate attributions. *Journal of Personality and Social Psychology*, Vol. 98, No. 5, 702–720.

O'Connor, T. (1994). An evolutionary argument against naturalism? *Canadian Journal of Philosophy*, Vol. 24, no. 4.

Penner, M. (2018). Cognitive science of religion, theism, and atheism. *Faith and Philosophy*, Volume 35, no. 1, January.

Petrovich, O. (1997). Understanding of non-natural causality in children and adults: A case against artificialism. *Psyche and Geloof*, 8, 151–165.

Piaget, J. (1972). The child's conception of the world. Towota, NJ: Littlefield Adams.

Plantinga, A. (1986). The foundations of theism: A reply. *Faith and Philosophy*, Vol. 3 No.3 July.

Plantinga, A. (1993). *Warrant and proper function*. Oxford: Oxford University Press.

Plantinga, A. (1994). *Naturalism defeated*. (unpublished manuscript). Retrieved from https://www.calvin.edu/academic/philosophy/virtual_library/articles/plantinga_alvin/naturalism_defeated.pdf

Plantinga, A. (1996). Respondeo, In Kvanvig, J. (ed.) (1996) *Warrant in contemporary epistemology: Essays in honor of Plantinga's theory of knowledge*. Latham, MD: Rowman & Littlefield Publishers.

Plantinga, A. (2000). *Warranted Christian Belief*. New York: Oxford University Press.

Plantinga, A. (2002a). Introduction. In Beilby, James (ed.). *Naturalism defeated? Essays on Plantinga's evolutionary argument against naturalism*. Ithaca, NY: Cornell University Press.

Plantinga, A. (2002b). Reply to Beilby's cohorts. In Beilby, J. (ed.). *Naturalism defeated? Essays on Plantinga's evolutionary argument against naturalism*. Ithaca, NY: Cornell University Press.

Plantinga, A. (2010a). Self-Profile, In Dancy, J., Sosa, E., Steup, M. (eds.), *A Companion to epistemology* (Second Edition). Oxford: Wiley-Blackwell.

Plantinga, A. (2010b). Reformed epistemology. In *A companion to philosophy of religion* (Second Edition). Taliaferro, C., Draper, P., & Quinn, P. (eds). Oxford: Wiley-Blackwell.

Plantinga, A. (2011). Content and natural selection, *Philosophy and Phenomenological Research* Vol. LXXXIII No. 2, September.

Plantinga, A. (2012). *Where the conflict really lies: Science, religion, and naturalism*. Oxford, UK: Oxford University Press.

Plantinga, A. (2015). *Knowledge and Christian belief*. Grand Rapid, MI: Eerdmans Publishing.

Plantinga, A., & Tooley, M. (2008). *Knowledge of God*. Oxford: Blackwell.

Pollock, J. (1986). *Contemporary Theories of Knowledge*. Savage, MD: Rowman and Littlefield.

Popper, K., & Eccles, J. (1977). The self and its brain. New York: Sringer International.

Poston, T. (2008). Internalism and externalism in epistemology. *Internet Encyclopedia of Philosophy.* Retrieved from https://www.iep.utm.edu/int-ext/#SH3c

Poston, T. (2014). *Reason and explanation: A defense of explanatory coherentism*. Basingstoke, UK: Palgrave-Macmillan.

Pryor, J. (2000). The skeptic and the dogmatist. *Nous,* 34, 517–549.

Pust, J. (2017). Intuition. *The Stanford Encyclopedia of Philosophy*, Zalta, E. (ed.). Retrieved from https://plato.stanford.edu/archives/sum2017/entries/intuition/.

Raman, L. & Winer, G. (2004). Evidence of more immanent justice responding in adults than children: A challenge to traditional developmental theories. *British Journal of Developmental Psychology*, 22, 255–274.

Reid, T. (1785/2002). *Essays on the intellectual powers of man*. Brookes, D. (ed.), University Park, PA: Pennsylvania State University Press.

Reid, T. (1788/2010). *Essays on the active powers of man*. Haakonssen, K., & Harris, J. (eds.), University Park, PA: Pennsylvania State University Press.

Rottman, J., Kelemen, D. (2012). Is there such a thing as a Christian child? Evidence of religious beliefs in early childhood, In: McNamara, P., & Wildman, W. (Eds.). Science and the world's religions: Origins and destinies. Santa Barbara, CA: Praeger Press.

Rottman, J., Zhu, L., Wang, W., & Seston, R., Clark, K., Kelemen, D. (2016). Cultural influences on the teleological stance: Evidence from China. *Religion, Brain and Behavio,* Volume 7, Issue 1.

Rysiew, P. (2011). Making it evident: Evidence and evidentness, justification, and belief. In Dougherty, T. (ed.), *Evidentialism and its discontents*. Oxford, UK: Oxford University Press.

Saxe, R., Tenenbaum, J, & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, 16(12), 995-1001.

Saxe, R., Tzelnic, T, & Carey, S. (2007). Knowing who dunnit: Infants identify the causal agent in an unseen causal interaction. *Developmental Psychology*, 43(1), 149-158.

Schwitzgebel, E. (2002). A phenomenal, dispositional account of belief. *Nous*, 36: 249-275.

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, 124(2), 209–215.

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the national academy of sciences*, 105(2), 809-813.

Simons, D. & Frank, K. (1995). An abstract to concrete shift in the development of biological thought: The insides story. *Cognition* 56, 129-163.

Sosa, E. (2007). *A virtue epistemology*. Oxford, UK: Oxford University Press.

Sosa, E. (1998). Minimal intuition. In DePaul, M. & Ramsey, W. (eds.), *Rethinking intuition*. Lanham, MD: Rowman & Littlefield Publishers.

Sterelny, K. (2006). Review of Daniel Dennett's Breaking the spell. *American Scientist*, September-October.

Sudduth, M. (1999). The Internalist character and the evidentialist implications of plantingian defeaters. *International Journal for the Philosophy of Religion*, 45, 167-187.

Sudduth, M. (2008). Defeaters in epistemology, *Internet encyclopedia of philosophy*. Retrieved from http://www.iep.utm.edu/ep-defea/

Swain, M. (1996). Warrant versus indefeasible justification. In Kvanvig, Jonathan (ed.) (1996) *Warrant in contemporary epistemology: Essays in honor of Plantinga's theory of knowledge*. Lanham, MD: Rowman & Littlefield Publishers.

Swinburne, R. (1979). *The existence of God*. Oxford, UK: Oxford University Press.

Swinburne, R. (2001). *Epistemic justification*. Oxford, UK: Oxford University Press.

Swinburne, R. (2010). Evidentialism. In A companion to philosophy of religion (Second Edition). Taliaferro, C. Draper, P., & Quinn, P. (eds.). Oxford, UK: Wiley-Blackwell.

Swinburne, R. (2011). Evidence. In Dougherty, T. (ed.), *Evidentialism and its discontents*. Oxford, UK: Oxford University Press.

Swinburne, R. (2016). *The coherence of theism* (second edition). Oxford, UK: Oxford University Press.

Swinburne, R. (2018). Phenomenal conservatism and religious experience, In Benton, M., Hawthorne, J. & Rabinowitz, D. (eds.), *Knowledge, belief, and God: New insights in religious epistemology*. Oxford, UK: Oxford University Press.

Talbott, W. (2008). Bayesian epistemology. *The Stanford Encyclopedia of Philosophy*, Zalta, E. (ed.). Retrieved from https://plato.stanford.edu/entries/epistemology-bayesian/

Talbott, W. (2002). The illusion of defeat, in James Beilby (ed.), *Naturalism Defeated?* Ithaca, NY: Cornell University Press.

Tang, S., Shepherd, S., & Kay, A. (2012). Do difficult decisions motivate belief in fate? A test in the context of the 2012 U.S. presidential election. *Psychological Science*, Vol. 25(4), 1046–1048.

Thurrow, J. (2013). Does cognitive science show belief in God to be irrational? The epistemic consequences of the cognitive science of religion. *International Journal for the Philosophy of Religion*, 74, 77-98.

Thurrow, J. (2014a). Does the scientific study of religion cast doubt on theistic belief? In Bergmann, M. & Kain, P. (eds.), *Challenges to moral and religious belief*. Oxford, UK: Oxford University Press.

Thurrow, J. (2014b). Some reflections on cognitive science, doubt, and religious belief. In Justin Barrett & Roger Trigg (eds.), *The Roots of Religion*. Farnham, UK: Ashgate.

Thurrow, J. (2016). Review of "A natural history of natural theology" (de Cruz and de Smedt). *Faith and Philosophy.* Volume 33, Number 3, July.

Tolhurst, W. (1998). Seemings. *American Philosophical Quarterly*, 35, 293-302.

Tooley, M. (2013). Michael Huemer and the principle of phenomenal conservatism. In Chris Tucker (ed.) *Seemings and justification: New essays on dogmatism and phenomenal conservatism*. Oxford, UK: Oxford University Press.

Tucker, C. (2010). Why open-minded people should be dogmatists. *Philosophical Perspectives*, 24(1), 529 – 545.

Tucker, C. (2011). Phenomenal conservatism and evidentialism in religious epistemology. In Kelly James Clark and Raymond J. VanArragon (eds.), Evidence and religious belief. Oxford, UK: Oxford University Press.

Tucker, C. (2013). Introduction. In Chris Tucker (ed.) *Seemings and justification: New essays on dogmatism*. Oxford, UK: Oxford University Press.

Turri, J. (2009). The ontology of epistemic reason, *Nous*, 43:3, 490–512.

Vahid, H. (2004). Varieties of epistemic conservatism. *Synthese*, Volume 141, Issue 1, pp. 97–122.

Vallortigara, G., Regolin, L, & Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biol* 3(7), e208.

van Woudenberg, R. (2013). Thomas Reid between externalism and internalism. *Journal of the History of Philosophy*, 51 (1), 75-92.

Vogel, J. (1990). Cartesian skepticism and inference to the best explanation. *Journal of Philosophy*, 87 (11), 658-666.

Wedgwood, R. (2001). Internalism Explained. *Philosophy and Phenomenological Research*, 65, 349–69.

Werner, P. (2014). Seemings: still dispositions to believe. *Synthese,* 191 (8), 1-14.

White, R. (2006). Problems for dogmatism. *Philosophical Studies*, 131 (3), 525-557.

Willard, D. (2004). The craftiness of Christ. In *Mel Gibson's 'passion' and philosophy: The cross, the questions, the controversy*. Garcia, J. (Ed.). LaSalle, IL: Open Court Publisher.

Williamson, T. (2000). *Knowledge and its limits.* Oxford, UK: Oxford University Press.

Williamson, T. (2007). *The philosophy of philosophy*. New York, NY: Routledge.

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128.

Pontifícia Universidade Católica do Rio Grande do Sul
Pró-Reitoria de Graduação
Av. Ipiranga, 6681 - Prédio 1 - 3º. andar
Porto Alegre - RS - Brasil
Fone: (51) 3320-3500 - Fax: (51) 3339-1564
E-mail: prograd@pucrs.br
Site: www.pucrs.br