Doctoral theses at NTNU, 2019:7

Halvor Kvandal

Prone to Believe in God

The Cognitive Science of Religion and its Normative Implications for Theist Religion

NTNU
Norwegian University of
Science and Technology

Doctoral thesis

Halvor Kvandal

ISBN 978-82-326-3624-2 (printed ver.)
ISBN 978-82-326-3625-9 (electronic ver.)
ISSN 1503-8181

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Humanities
Department of Philosophy and Religious Studies

Doctoral theses at NTNU, 2019:7

NTNU
Norwegian University of
Science and Technology

Halvor Kvandal

Prone to Believe in God

The Cognitive Science of Religion and its Normative Implications for Theist Religion

Thesis for the Degree of Philosophiae Doctor

Trondheim, November 2019

Norwegian University of Science and Technology
Faculty of Humanities
Department of Philosophy and Religious Studies

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Humanities
Department of Philosophy and Religious Studies

ISBN 978-82-326-3624-2 (printed ver.)
ISBN 978-82-326-3625-9 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2019:7

Printed by NTNU Grafisk senter

Prone to Believe in God

The Cognitive Science of Religion and its Normative Implications for Theist Religion

by

Halvor Kvandal

Acknowledgments

My mother Lilleba Lund Kvandal (1940-2016) has been a great source of inspiration. Our conversations about topics like Carl Gustav Jung, Norwegian folklore, mysticism, and Christianity provided perspectives and thoughts that gave me an interest in examining religious issues from a philosophical perspective.

I am deeply grateful to my supervisors Jonathan Knowles and Gabriel Levy for all their help and support, which has been crucial for me in writing this thesis. I would also like to thank Aku Visala, John Teehan, Istvan Czachesz, Atle Ottesen Søvik, Robert L. McCauley, Jan-Olav Henriksen, Asbjørn Dyrendal, Ingebjørg Seip, Lari Launonen, Mons Andreas Nyquist, Morten Langfeldt Dahlback, Espen Stabell, Sigurd Arnold Enger, Jørgen Wold Akselberg, Pål Røland, and Helen De Cruz.

I would like to thank the Department of Philosophy and Religious Studies at NTNU and Kari B. Berg in the administration.

Finally, I want to thank my beloved wife Claudia Vargas, who has supported, inspired, and encouraged me so much throughout this process.

Contents

Chapter 1. The place of Cognitive Science in the Science and Religion Debate

1. Unweaving the rainbow and unweaving religion

Did Newton's finding that the rainbow consists of normal white light refracted into its component colours by drops of water take away its poetic beauty? The biologist Richard Dawkins (1998) opens a discussion about the relationship between science and poetry with this question. He uses a poem by John Keats about how "cold philosophy" makes nature "dull" and "ordinary" to illustrate the view that by providing a scientific explanation of natural phenomena, we take away their poetic beauty or sense of mystery. In response, Dawkins suggests that science does not remove the poetic beauty of nature. Rather, the opposite is the case. By "unweaving the rainbow", Newton provided insights of even greater poetic quality (Dawkins, 1998, 38-43). Science is thus not only compatible with poetry but also conducive to it.

The beliefs that natural phenomena are purposefully brought about by supernatural agents or that the world is populated by such beings seem to be obstacles to the scientifically informed poetry Dawkins is describing. Then, while science and *poetry* can come together in a fruitful meeting, its relationship with *religion*, which contains these beliefs, seems conflictual. Does science help establish the falsity of supernatural belief? In some cases, yes. Consider the belief that the rainbow can *only* arise when the gods decide to create it. By creating a rainbow with a prism, Newton demonstrated that this is wrong. It is less clear that science establishes the falsity of supernatural beliefs in general. For example, one cannot use a prism-like experiment to demonstrate that there are no supernatural agents or no divine or transcendent reality. Tensions nevertheless sometimes manifest between science and religion. These become especially vivid when scientific progress occurs in contexts where religion is seen as providing comprehensive answers, also to questions about natural phenomena. Tensions sometimes worsen into full-blown conflicts with political and legal ramifications. An example is the conflict between Christian fundamentalists and the US educational system about evolution in the last century (Larson, 2003). In many cases, tensions and conflicts between scientific theories and religious ideas or doctrines are resolved through reinterpretations and revisions of the latter.

Religious ideas, or at least those that non-fundamentalists promote, are often malleable and therefore potentially able to incorporate scientific findings that initially seem troubling. They thus add a supernatural gloss on science, while abstaining from making substantive claims about scientific issues. All things considered, how conflictual is the relationship between science and religion? We will not approach such a broadly construed question head on but focus on a particular meeting between science and religion. Here, science and religion do not meet in virtue of making competing claims about the same natural phenomena. They meet in virtue of how the former tries to explain the latter.

Various areas of science have during the recent three decades come together in an attempt to "unweave" (to use Dawkins' phrase) a complex and unwieldy phenomenon, namely religion itself. While Newton showed that the rainbow consists of *ordinary white light*, many scholars engaged in this attempt argue that religion is the by-product of *ordinary cognitive processes*. Dawkins (1998) denies that Newton belittled the rainbow in virtue of explaining it, but he appeals to this unweaving of religion in his (2007) attack on religion. He therefore seems to view the prospect of explaining religion as different from that of explaining other phenomena. Philosophers and theologians have also recently discussed what implications this science of religion has for religion. They have sensed a worry. What kind of worry is this? Consider a related discussion about the prospect of explaining morality. As Guy Kahane points out, scholars feel uneasy about the possible implications an evolutionary study of human morality has for the status of evaluative beliefs. This uneasiness has to do with how different morality seems from the immersed subjective perspective of moral agents and how it seems from the detached perspective of evolutionary theory (Kahane, 2011, 104). In extension, it also relates to the possibility that we can explain morality without presupposing that our moral beliefs are true. The worry about the prospect of explaining religion seems to involve similar considerations. For example, also religion seems quite different from the perspective of those who believe and from the perspective of empirical scholarship. Furthermore, it now seems possible to explain the phenomenon of religious belief without presupposing that any religious claim is true.

In considering this, a barrage of questions faces us. Does the uneasiness many feel about explaining religion indicate the presence of an underlying conflict between science and religion? Is Dawkins right to exempt religion from phenomena that science does not belittle in virtue of explaining? Do explanations of religion have any impact on the status of religious beliefs or doctrines? If so, is it positive or negative? Can a science of religion inspire poetry about religious ideas in a way akin to how a science of the rainbow inspires poetry about the

refraction of light? Can religious believers plausibly integrate a scientific account of religion into their religious worldview? Before we look closer at these issues and what the present investigation will do, let us consider the science of religion with which we will be concerned.

The thesis that religion is the by-product of ordinary cognitive processes is found in an area known as the *Cognitive Science of Religion* (CSR).[1] By this, I mean a loosely delineated area in the current study of religion that combines cognitive psychology, evolutionary theory, and anthropology in an attempt to explain religion as a natural phenomenon. Scholars that have contributed in this area include, among others, Stewart Guthrie, E. Thomas Lawson, Robert N. McCauley, Deborah Kelemen, Harvey Whitehouse, Pascal Boyer, Scott Atran, Justin Barrett, Lee Kirkpatrick, Richard Sosis, Joseph Bulbulia, Helen De Cruz, Ara Norenzayan, Claire White, and Aiyana Willard.

The anthropologist Pascal Boyer (1994, 2002, 2018) views religion as the phenomenon whereby a certain range of concepts involving purported intentional beings with counterintuitive properties tends to propagate and thereby gain stability in human cultures over time. Concepts of gods, spirits, ghosts, and ancestors tend to achieve this because the way our mind works when processing information in general also makes it susceptible to obtain and transmit such concepts. This example illustrates two central features of this kind of research. The first is a focus on *cognition*, the processing of information. The study of how the mind does this plays an important role in explaining religion. Another central feature is a focus on *evolution*. To claim that religion is a by-product is to claim that it is a by-product of something else, namely adaptations. Adaptations are complex functional structures, traits, or behaviors that have evolved as the result of how they increase survival and reproduction (i.e. fitness) compared with other variants (Atran, 2002, 22-23). An example of an adaptation is the human umbilical cord, which has the belly button as a by-product (Buss, 2015, 36).

By-product theories of religion do not focus directly on the evolution of behaviors but deal with alleged underlying cognitive systems and proximate processes involved in how we think and act. To be more specific, they focus on alleged *psychological adaptations* of relatively recent origin. *Evolutionary psychology* studies such adaptations and provides a backdrop for this approach. The picture we are presented with is something like the following. In our ancestral past, successful handling of challenges related to foraging, hunting, predator-avoidance, mating, social exchange, attachment, coalition building, and so on, was crucial for survival and reproduction. Psychological adaptations evolved as specialized solutions to these

[1] The main alternative to the by-product view is an *adaptationist view of religion.* Chapter 2, section 2 discusses the by-product versus adaptationionism issue.

and other challenges. The functioning of these adaptations influences how we think and act today, because we are descendants of those who were the most successful in handling them. The primary period in which many of these adaptations evolved is roughly something like the Pleistocene, where hominids lived largely as hunters and gatherers (Cosmides and Tooby, 1992, Buss, 2015).[2] The Pleistocene is the period from roughly 1.8 million years ago until the agricultural revolution approximately 12 000 years ago. Homo sapiens is the only surviving hominid species that appeared during this period (Boyd and Silk, 2015, 212).

Pascal Boyer and the psychologists Justin Barrett and Lee Kirkpatrick, building on evolutionary psychology, see our susceptibility to belief in deities, spirits, demons, and so on, as a by-product of the normal functioning of psychological adaptations specialized to deal with the aforementioned challenges, such as those related to detecting agents and handling hominid social interaction. The anthropologist Stewart Guthrie also appeals to evolutionary considerations in his work on religion. In his view, the evolution of the hominid mind has favoured a perceptual strategy featuring a low threshold for inferring the presence of humanlike agents. This makes us susceptible to over-attribute humanlike agency to the natural world. In virtue of thus anthropomorphising the natural world, we have given rise to religion (see Guthrie, 1980, 1995, 2017). In other words, the notions of spirits and gods found in religion are ultimately our creation, brought about as the unintended result of an adaptive strategy in hominid agency-detection. Pascal Boyer and Justin Barrett focus on the structure of religious concepts. They argue that those concepts we can expect to propagate have a peculiar structure: they largely agree with but to some extent also violate intuitive expectations we have towards the world. This structure is *cognitively attractive*. This attractiveness consists of catchiness, memorability, and the ability to easily become embedded in intuitive thought. Cognitive attractiveness is central to why the aforementioned concepts increase their frequency in human populations and remain widespread over time. Religion is thus explainable in terms of the underlying cognition it involves (See Boyer, 1994, 2002, 2018, Barrett, 2000, 2004, Kirkpatrick, 2005, see also Atran, 2002 and Sperber, 1996).

Will Gervais and Joseph Henrich (2010) build on the notion of religion as consisting of cognitively attractive agent-concepts but deny that it is sufficient to fully explain religion. They emphasize the importance of adaptations for cultural learning and the biases and heuristics they involve. To explain religion, we also need to deal with how subjects obtain and use information found in their specific social context. In this way, we can better explain something we all readily

[2] This thesis has been strongly criticized. See for example criticisms by Buller (2006) and Richardson (2007). H. C. Barrett (2015) responds to recent criticism of evolutionary psychology and defends an updated version.

observe: people in different cultures tend to believe in different deities (Gervais and Henrich, 2010, Gervais et al. 2011, see also H.C. Barrett, 2015). Cultural input is central in discussions of how religious beliefs arise (see e.g. Barrett and Lanman, 2008 and Van Leeuwen and Van Elk, 2018).

Now that we have a taste of the scholarship the next chapters explore in greater detail, we are ready to close in on the main research-question the present investigation asks. First, here is a brief outline of the remainder of the chapter. Section 2 presents the main research question. The subsequent three sections relate the present investigation to ongoing debates about the philosophical implications of the CSR. I use a distinction between three kinds of *normative implications*, to be specified below, as a map to navigate these debates. Section 6 outlines the remaining chapters.

2. Three normative implications and three related discussions

The main question the present investigation asks is the following:

What normative implications for theist religion do central theories and claims in CSR have?

This is not a sociological question about how the emergence of a particular kind of research affects religion. It is a normative question about how central theories, claims, and presuppositions in this area, *if true or strongly plausible*, normatively affect or bear on claims, arguments, and beliefs in what I call *theist religion* (a notion whose meaning I clarify below). The following statements describe what I take to be normative implications:

(1) The truth of claims in A bears logically on the truth of claims in B.
(2) Claims in A have a bearing on philosophical arguments or theses connected to B, and it is therefore relevant to consider the former when assessing the latter.
(3) Claims in A have a bearing on the epistemic status of beliefs in B, and it is therefore relevant to consider the former when evaluating the latter.

A has normative implications in sense (1) for B if A makes explicit claims or involve tacit presuppositions whose truth-value affects the truth-value of claims in B. One example of this is that the truth of central claims in B is less likely given the truth of explicit claims or tacit

presuppositions in A. Another example is that the truth of claims or tacit presuppositions in A entails the falsity of claims in B. Furthermore, if the truth-value of a claim that follows logically from claims or presuppositions in A affects the truth-value of claims in B, then A also has normative implications for B. Thus, if the existence of God is less likely given CSR or if the truth of such research entails that there is no God, then CSR has normative implications in sense (1) for theist religion. We will see that scholars make claims and presuppose things that have such implications if true.

A has normative implications for B in sense (2) if claims in A are relevant to consider when assessing arguments or theses connected to B. For example, if claims in CSR are relevant to consider when assessing an argument for or against the existence of God, or when assessing an argument or thesis about the justification (rationality, warrant, and so on) of theistic beliefs, then such research has implications for theist religion in this sense. The presence of such implications might entail implications in senses (1) and (3) too. For example, if claims in A bear on the plausibility of arguments for or against the truth of claims in B, then the former claims also bear logically on the truth of claims in B. Furthermore, if claims in A bear on arguments on which beliefs in B are based, then the former claims bear on the epistemic status of these beliefs.

A has normative implications for B in sense (3) if A consists of claims that have relevance in an evaluation of the epistemic status of beliefs in B. Claims that have this kind of relevance say something about factors that affect the epistemic status of beliefs. CSR has normative implications for theist religion in this sense if it makes claims that relevantly bear on the justification, virtue, rationality, or warrant of such beliefs. Suppose theistic beliefs are based on arguments for the existence of God. As mentioned, if CSR bears on these arguments, then it bears also on these beliefs. However, this is not the only way for such research to have implications in sense (3). I shall argue that in virtue of explaining beliefs, we can also by extension evaluate them. The central premise here is that the etiology of beliefs is epistemically crucial. This means that if CSR sheds light on the etiology of theistic beliefs, then it has implications in sense (3) for such beliefs, whether that etiology is a careful reasoning process involving arguments or a largely non-conscious and intuitive process.

The notion of "theist religion" is not intended as a definition of religion but as a shorthand for three things. The first is *theism*, the proposition that there is a perfectly good, omnipotent, omniscient, personal being who has created the universe and everything else which exists.[3] By

[3] This proposition is sometimes referred to as *classical theism.* Classical theism is at play in central contributions to analytical philosophy of religion, such as those we shall discuss, namely the work of Alvin Plantinga (2000)

"God", I mean the purported culturally specific deity theism describes, which according to central tenets in Judaism, Christianity, and Islam has interacted in important ways with purported historical figures like Abraham, Moses, Jesus, and Muhammad. I use "God" only for this deity and therefore not for Zeus, Thor, the Feathered Serpent, or the Mayan deity of Maize, which are purported *gods or deities* of other cultures. Second, in "theist religion", I include philosophical arguments and theses tied to theism. Arguments and theses tied to theism are about this proposition. For example, consider arguments for the conclusion that theism is true or theses according to which theism is compatible with scientific findings. We find these in the intellectual branches of Abrahamic theistic traditions, that is, in their theology and philosophy. We can relate these two elements of "theist religion" to implications in senses (1) and (2). Third, in "theist religion", I include beliefs about God which we plausibly can attribute to subjects. I call these "theistic beliefs".[4] We can relate theistic beliefs to implications in sense (3).

Those who see CSR-scholarship as problematic or misguided in general will of course object to the notion that findings in this area have important normative implications.[5] Normative implications of claims that emerge from a problematic research paradigm are not obviously significant. In my opinion, CSR is not wholesale problematic. The central idea that the study of cognition and evolution sheds important light on religion is for example plausible. Moreover, even though many of the details are objects of intense debate (as should be the case in a healthy scientific field), it seems that there is empirical support for central claims. However, it would go beyond the boundaries of the present study to assess the viability of this field as a whole. For those who are sceptical, I can be seen as investigating normative implications of CSR conditional on the truth or plausibility of its central theories and presuppositions. I shall assume that CSR is a legitimate research-enterprise that can face down important objections, and that

---

and Richard Swinburne (2004). For definitions of God along the lines of classical theism, see Plantinga (2000, vii) and Swinburne (2004, 93). See also the overviews in Taliaferro (1998) and Peterson et al. (1997) for examples of the centrality of this notion of God in philosophy of religion. To indicate God in this sense, the text uses capital "G". I treat the philosophical arguments and theses we encounter in chapter 4 and 5 as concerned with God in this sense. But since the present investigation also engages thoroughly with the psychological study of religion, a potential ambiguity arises. We must be aware that scholars engaged in this study often use "God" (with or without capital "G"), "gods", or "deity" in much broader ways. For example, Ara Norenzayan employs the term "Big God" for, "omniscient, all-powerful, morally concerned deities who directly observe, reward, and punish social behaviour" (Norenzayan, 2013, 127-128). "Big God" has a broader meaning than "God", since Norenzayan for example says that Big Gods do not necessarily denote the idea of an only God or a creator (Norenzayan, 2013, 130). As another example, Justin Barrett and Kelly James Clark use "gods" to denote, "any supernatural intentional agents whose existence would impinge upon human activity" (Clark and Barrett, 2011, 652). This is an even broader conception where ghosts, spirits, angels, or demons count as gods. To avoid confusing these terms with what I call "God", I shall when possible use the specific terms scholars employ. But when scholars use "God" in ways that differ from mine or where it is not clear that their usage is the same, I indicate this with footnotes.

[4] This means I do not use "theistic beliefs" for beliefs about other deities than God.

[5] For criticisms of this field, see Bloch (2002), McCutcheon (2010), and Jones (2016).

its central theories are *approximately true or at least strongly plausible*. I shall present the selection of work I find the most compelling and defend it from some recent criticism. But ultimately, I leave it for scholars in this area to convince us of the truth of central theories. The claims and theses currently available are sufficiently bold and interesting to make their implications, *if true*, worth investigating philosophically. That is our main task.

Let us now turn to current discussions about what philosophical implications CSR has. The three normative implications specified above provide a way to navigate these. We can loosely relate normative implication in sense (1) to discussions of whether CSR and theism are logically compatible. We relate normative implication in sense (2) to debates about the relevance CSR has in philosophy of religion. Finally, we can relate normative implication in sense (3) to debates about the epistemic relevance of CSR for theistic beliefs. These three debates all relate to the larger overarching debate about how science and religion relate. The next section looks closer at implications in sense (1) and provides some introductory remarks about the science and religion debate.

3. Is the cognitive science approach compatible with theism?

Can central claims and presuppositions in CSR and theism both be true or are they logically incompatible? An investigation of this sheds light on one way in which research in this field might have implications for theist religion. As we shall see, some scholars make claims that suggest conflict, but it remains an open question whether this conflict is resolvable. Interestingly, CSR seems to be in a more serious conflict with something other than theism, namely a particular kind of attitude towards the study of religion which CSR, if viable, debunks.

Some discussions of how science and religion relate deal with what theoretical model or description best captures their relationship. These models are usually *global*, i.e. attempts to apply one metaphor or picture to capture how all of science and all of religion relate. We can distinguish the question of what a right *global model* is from what a right *local model* is. The latter applies to the relationship between a particular form of science and a particular form of religion. The question whether CSR is compatible with theism relates to the issue of what a correct *local* model for *CSR* and *religions committed to theism* looks like. Answering it can also tell us something about how a correct global model looks because a good way to provide a

correct global model is to generalize from outcomes of specific encounters between science and religion.

Here are some models of how science and religion relate. According to "independence", science and religion are simply two different things that have little if anything to do with each other. A central claim related to this model is that religion is not in the business of making factual claims about natural reality. Gould (2002) for example argues that science and religion are two domains of human activity that do not overlap. Science concerns the factual whilst religion is solely about meaning and purpose. This independence-model entails that the two are compatible, because there are no claims in religion which could be at odds with scientific claims.[6] "Conflict" sees science and religion as in principle incompatible or at least as being at odds in serious ways. Advocates of this model view science and religion as similar in the respect that both involve claims about reality. They then argue, for example, that these claims contradict each other, that science and religion represent incompatible methods for obtaining knowledge, or that a scientific attitude is incompatible with attitudes that are central or constitutive for religion.[7] Scholars have also proposed models such as "dialogue", "integration", and "reconciliation" (see Stenmark, 2010, 278-295, cf. Ruse, 2011). Reconciliation presupposes that there are tensions and possible conflicts between science and religion but claims that the two are possible to reconcile in such a fashion that the distinct and unique features of each are preserved (Stenmark, 2010, 279). As Mikael Stenmark points out, we can envisage two quite different views about conflict. The first sees the conflict as irresolvable: science and religion are in principle impossible to reconcile. This relates to the conflict model. The other view,

[6] Peterson et al. (1997) describe two further perspectives one can see as supportive of the independence-model. First, from the perspective of the protestant theology of Karl Barth, one might argue that science and religion are two entirely different things, which do not come into contact with each other. Religion (in this case Christianity) is about how a transcendent God revealed himself in Jesus. The goal of religion is that we have a personal relationship to God. In contrast, science is about obtaining knowledge of the natural world. This is an entirely different business, which need not have anything to do with our relationship to God. Second, from the perspective of a philosophy of language based on the later Wittgenstein, some have argued that science and religion are two different "language-games", which are characterized by their own distinct logic. Based on this, they contend that something like an independence-model is correct (Peterson et al. 1997, 224-226).

[7] A central assumption in this model is that science and religion compete in the same business, and that the latter must meet the standards of the former. For example, John Worrall (2004, 64) argues that one cannot be properly scientific and a religious believer at the same time. All claims must meet the standards of science. Religious claims fail to meet such standards, since they are "taken on faith" and fail to yield testable predictions. One therefore cannot accept them without violating a proper scientific attitude (Worrall, 2004, 59-68). Jerry Coyne (2015) argues in a similar way, suggesting that both science and religion attempt to obtain knowledge of reality and therefore must be assessed as competing enterprises. These enterprises are incompatible because of conflicts between their methods and their respective philosophies. For example, science involves critical scrutiny of evidence while religion involves taking beliefs on faith (Coyne, 2015, 63-96). Pascal Boyer (2002) is critical of those who make global claims about something we call "science" and "religion". He does not accept that these are coherent wholes with essential natures. Boyer nevertheless contends that "science" shows us that "religion" is fundamentally *defective* as a source of knowledge (2002, 368-369), thus chiming in with these conflict-advocates.

which we find in the reconciliation-model, grants that there are tensions and conflicts but sees them as resolvable, thus maintaining that science and religion can be reconciled (Stenmark, 2010, 281).

We will now see examples of how scholars in CSR make claims that are seriously at odds with theism. This indicates that there is conflict, but it does not thereby establish that the conflict-model is correct. It is still an open question whether the two can be reconciled. It is initially surprising that there are such tensions. Scholars in CSR often describe this area as committed to *methodological naturalism* (MN). This principle requires a bracketing of evaluative and metaphysical issues. For example, a scientist should not insert evaluative judgements or religious (or atheist) beliefs into his scientific work on religion. The goal is to explain belief in such things, not to investigate whether they are real (Kirkpatrick, 2005, 5-8).

Not everyone in CSR agrees. Here then is our first example. Edward Slingerland and Joseph Bulbulia (2011, 308-312) propose that metaphysical claims about the existence of immaterial minds are not acceptable because they are not empirically defensible. CSR should therefore remove the brackets around the existence of immaterial minds and dismiss the notion that there are such things. What normative implications for theism does this proposal have? Let us propose an argument for illustrative purposes:

P1: Science should dismiss the existence of immaterial minds on empirical grounds.
P2: The God of monotheist religion is an immaterial mind.
C: Science should dismiss the existence of God on empirical grounds.

Theism is the thesis that an immaterial personal being exists. For example, William Lane Craig and J. P. Moreland (2009, x) contend that theism involves the claim that an immaterial divine *mind* exists. Richard Swinburne, a central theist philosopher whose work we consider later, describes theism as the thesis that a non-embodied (in other words immaterial) spirit or person exists (Swinburne, 2004, 7, 93, cf. Plantinga, 2000). If CSR should deny the existence of immaterial minds and God is such a mind, then this area of science should deny the existence of God. If we accept Slingerland and Bulbulia's recommendations as representative for CSR, then something like what the theologian Aku Visala calls the "falsity of religious beliefs thesis" is true. According to this thesis, CSR is incompatible with propositions theists cannot give up without giving up theism (Visala, 2011, 158-159).[8] We should appreciate that the

[8] The proposal that CSR denies the existence of immaterial minds also has severe implications for natural theology. According to Craig and Moreland (2009, x), the notion of an immaterial mind is *constitutive* of natural theology,

incompatibility with the belief in an immaterial mind seems much more serious than for example an incompatibility with creationist thought. This is because the notion that God is immaterial is central to theism in a way creationism is not.[9] If the denial of the existence of immaterial minds is integral to CSR and God's immateriality is integral to theism, then we do not achieve a reconciliation by removing such elements from them. This is because we then do not preserve the distinct features of such research or theism but change them into something else (see Stenmark, 2010). But does this mean John Worall (2004) is right that a choice *must* be made, in this case, between accepting CSR or remaining a theist? Many will deny this by denying that CSR should remove the brackets around the existence of the supernatural the way Slingerland and Bulbulia propose. In other words, they will argue that CSR should remain committed to methodological naturalism.

Consider next the work of Stewart Guthrie, a central scholar in CSR who claims that religion is *anthropomorphism*, an attribution of human features to non-human phenomena. Guthrie views this attribution as *the making of a mistake* (1995, 62-64, 200). Then, if religion is anthropomorphism, religion is a mistake. Guthrie makes this explicit:

> "If religion is generalized and systematic anthropomorphism, can it be said simply to be a mistake? If it is only a mistake, why does it persist? These questions may seem presumptuous but they do not seem unanswerable. Nor are they irrelevant to humanistic, or even to social-scientific, concerns. On the contrary, they bear directly on our understanding of human thought and action. My answer is brief. The central religious assertion, that the nonhuman world is, in whole or in part, significantly humanlike, seems mistaken." (Guthrie, 1995, 200).

---

and to dismiss that claim is therefore "abortive" for that intellectual enterprise. If CSR contains this proposal, then such research has implications in sense (2) of the strongest kind: it shows us that natural theology must be aborted.

[9]With creationism, we can understand a variety of religious anti-evolutionary views about the origin of humans (Blancke et al. 2014, 5, Numbers, 2011 provides detailed historical overview). Many of these emerged in the US in the 20th century, where organized anti-evolutionary movements grew up. The first version that came to the forefront was *old earth creationism*, which holds that the earth is very ancient and that human life is the very recent product of a divine act. In the latter half of the 20th century, *young earth creationism* emerged. This is the thesis that God created the universe, the earth, and human beings as recently as 10 000 years ago (Blancke et al. 2014, 2-5). Polls since the early 1980s have repeatedly found that almost half of the US population supports some sort of creationism, while a somewhat smaller proportion accepts the notion that God guides the evolutionary process. Only a minority (between 9 and 16 %), believe that humans are the product of an unguided evolutionary process (Numbers, 2014). Though creationism is widely denied by theologians, it is therefore not a peripheral religious idea. The relative prominence of creationist thought means that the truth of CSR, which presupposes evolutionary theory, entails the falsity of many theistic beliefs. But that does not mean it entails the falsity of theism. Theists for example open up for the notion of a divinely guided evolution (see e.g. Paul II, 1997, Haught, 2011, Alexander, 2014). We will nevertheless see (chapters 4 and 5) that findings in CSR can create problems for that sort of view.

It seems to follow from the truth of this thesis that theism is a mistaken claim. This is because theism places humanlike agency at the heart of reality: the universe and everything within it has been brought about by an invisible intentional agent with whom we can communicate. Some theists would also say that God is mysteriously present in the universe and the natural world. Then obviously, large parts of non-human reality contain agency. But if it involves anthropomorphism to believe that this is true, and anthropomorphism by definition is a mistake, then theism is not correct.[10] This example indicates that CSR research, at least to the extent it commits to Guthrie's thesis, has strong normative implications in sense (1) for theist religion.

Let us look at a final example of this type. We saw above that if the truth of a presupposition tacitly endorsed by CSR entails the falsity of theism then this also establishes normative implications in sense (1). Is there evidence of such a presupposition? Aku Visala argues that there is. CSR tacitly tends to presuppose something he calls "strict naturalism". This thesis includes physicalism, the view that all that exists is physical. Strict naturalism is according to Visala a strongly atheistic thesis. He further suggests that if CSR incorporates it, then such research is in direct conflict with theism and many other religious propositions (Visala, 2011, 86-93). Here we have an example of strong normative implications in sense (1) for theist religion: the truth of a presupposition in CSR entails the falsity of theism. Visala argues in relation to this and similar examples that we should re-interpret CSR. We should reject physicalism and relate what is alleged to be a more plausible philosophical framework to this enterprise, which is less hostile to theism. Visala (2011) proposes that what he calls "broad naturalism" can do this job. He argues that this thesis is philosophically plausible in its own right and that wedding it to CSR does not require any important changes in that kind of enterprise. This opens up for reconciliation. We shall not go further into this argument here, but will return to Visala's discussion of CSR in section 5 and again in chapter 3, section 2.

These three examples indicate that there are tensions or even full-blown conflict between CSR and theism, but they do not establish an irresolvable conflict. It seems possible, something Visala's proposal illustrates, to re-interpret CSR in such a way that the conflict is resolved. Let us briefly look at some challenges that attempts to reconcile science and religion in general face. Stenmark points out that some attempts to resolve tensions involve significant re-interpretations or alterations with different priorities. He asks us to consider two groups of people. The first suggests that we keep religion intact but revise evolutionary science in dramatic ways such that no conflict with religion obtains. The second group suggests we keep

[10] Conversely, if theism is correct, then the formation of theistic beliefs is not an example of anthropomorphism. This is because if God exists, then theists do not make a mistake when forming belief in God.

science intact but instead dramatically rethink religion. The former group gives religion priority and contends that science must change whilst the latter gives science priority and alleges that religion must change (Stenmark, 2010, 281-282).

Such attempts are problematic because of the extent to which they involve alterations (in religion or in science). A plausible reconciliation-attempt must be fair in its interpretation of the elements it tries to reconcile. Reconciliation requires, as Stenmark (2010) points out (mentioned above), that we preserve important features of what we reconcile. We do not plausibly achieve a reconciliation between religion and science if we make dramatic changes to how we conceptualize the two. For example, we saw that the independence-model entails that there is no conflict, but it hardly therefore reconciles a plausible construal of *religion* with science. This is because religion plausibly involves claims[11] and therefore is not only about meaning and value. One also fails to reconcile the two if one alters science in problematical ways, for example by removing mainstream evolutionary theory to achieve compatibility with creationism. If we remove or adjust evolutionary theory to accommodate creationist inclinations, then we have simply changed CSR into something else and no reconciliation is achieved. In contrast, Visala's removal of "strict naturalism" from CSR (see his 2011) plausibly does not involve such an alteration of science. For example, it does not seem necessary for CSR to commit to *physicalism*, a strong metaphysical thesis which is contentious both among theists and non-theistic philosophers (for overview and discussion, see Ritchie, 2008).

A plausible reconciliation-thesis not only abstains from making dramatic alterations but shows how the reconciliation of full-blooded versions of the two elements is possible. In other words, such a thesis does not relate anaemic or watered-down versions of CSR and theism. What makes theism striking is the boldness of its central assertion that there is a bodiless, omnipotent, omniscient, and perfectly good person who created the universe (see e.g. Swinburne, 2004). A feature which makes CSR striking is what we can call its explanatory greediness or boldness. Appreciating this feature of CSR is therefore relevant for those who attempt to reconcile it with theism.

[11] I shall assume for the purposes of the present investigation that religion (in a colloquial sense and not to be confused with my stipulative shorthand "theist religion") involves claims in the following two senses. First, in religions we find a discourse which includes elements we can interpret as utterances that express claims. Though we should not view *all* the utterances religious believers make or *all* the written material we find in a given religious tradition as expressing claims, we should do this with respect to at least *some* of its discourse. We can refer to this view of religious discourse as *cognitivism.* The main alternative to this view is *non-cognitivism.* According to non-cognitivism, religious discourse is not propositional (at all). An example is the view that religious utterances do not express factual claims but only communicate or give expression to attitudes or feelings without propositional content (see Moore and Scott 2007 for more). Second, I assume that religion involves claims in the sense that religion consists (among other things) of the variegated beliefs that people have. We can, for the purposes of the present investigation, understand beliefs as propositional attitudes (i.e. attitudes towards claims).

Those who *in addition* attempt to reconcile CSR-research with a set of influential, traditional views on the study of religion will arguably have a harder time than those who focus on the relationship this research has to theism. Let us digress to appreciate this, thus situating CSR alongside other approaches to the study of religion. These traditional views involve concerns that relate to the worry Keats gave expression to in his poem about the rainbow. Recall that Keats (the way Dawkins presents him) worries about how "cold philosophy" disenchants us with its mechanical understanding of the world, thus removing the sense of poetry that nature instils. This seems to be a worry that scholars of religion who object to the cognitive approach also share. Imagine what happens when these scholars encounter statements like the following:

> "{I} want to understand everything about religion: not just images of God, or religious attributions of causality, or spiritual experiences, or religious development, or prayer – all of it. I am intellectually greedy" (Kirkpatrick, 2005, 7).

We see a similar explanatorily bold attitude in the work of Pascal Boyer (2002) and Stewart Guthrie (1995). They provide ambitious explanatory theories of religion. It seems that explanatory greed is representative for CSR as a research-enterprise in general (Jones, 2016). Such research then represents a sharp break with *hermeneutical theories of religion*, which focus on the interpretation and understanding of religious texts, experiences, and behaviours, and sometimes even appeal to a mystical, religious feeling in its interpretations of religion. Wayne Proudfoot traces the legacy of these theories to the work of Friedrich Schleiermacher, an influential 19th century Christian thinker who set out to defend and preserve what he saw as the inherent uniqueness of religion from enlightenment criticism. Proudfoot further points out that hermeneutical scholarship of religion such as that found in the work of Rudolph Otto and Clifford Geertz continue this kind of approach, which sees it as in principle wrong to *explain* religion because this phenomenon carries meaning and therefore requires the application of hermeneutical methods only (Proudfoot, 1985, 52-55). The explanatory ambitiousness CSR exhibits is strongly at odds with such an approach, especially the thesis that hermeneutics is the *only* viable way to study religion (for further discussion, see McCauley, 2017a). We can hardly reconcile that thesis with CSR.

Is religion a mystical phenomenon some of us experience and which we at best can understand or interpret in a piecemeal fashion but never actually explain? Daniel Dennett (2006) argues against such an idea, contending that it is time to "break the spell" of religion. By the notion of a "spell", Dennett primarily means the ubiquity in society, public discourse,

and among religious believers of the idea that religion is in principle beyond the reach of science and that it is inherently different from other phenomena (Dennett, 2006, 17). Dennett's view on the need to break this spell chimes in very well with the prevailing attitude in CSR; that religion is something we both can and should explain. In extension, Dennett means something more with "breaking the spell". He points to something that *could* happen if we break the first spell: religion loses its ability to enchant or entrance believers and fascinated by-standers. This is according to Dennett akin to what happens when one is deeply moved by the music at a concert but then a phone rings, ruining the moment. In short: by breaking the first spell we might end up damaging what we try to explain, leaving it unable to entrance us the way it used to (Dennett, 2006, 12-18). The situation for religion then seems different than for the rainbow: while a science that explains the rainbow enables poetry to arise, a science that breaks the spell by explaining religion takes away the poetry. But *must* we tie the sense of poetry in religion to the thesis that it is inherently misguided to explain it? This is far from clear.

To return to our discussion of how CSR relates to theism, can theists plausibly accept the notion that religion is a phenomenon we can (at least in principle) fully explain naturalistically? Can they plausibly accept the view that we can and should "break the spell" (at least in the primary sense Dennett describes)? This is not the case if breaking the spell necessarily leads us down a path towards a physicalist thesis which brings about full-blown conflict with theism. But Dennett argues that this is not necessarily the case with respect to a naturalistic science of religion:

> "Notice that it could be true that God exists, that God is indeed the intelligent, conscious, loving creator of us all, and yet *still* religion itself, as a complex set of phenomena, is a perfectly natural phenomenon." (Dennett, 2006, 25).

This indicates that Dennett is not an advocate of the view that *logically* it is not possible for CSR and theism both to be true. If there is a way to handle the tensions we saw examples of above, then he seems right. What about the thesis that *religion is a natural phenomenon*? Is that something theists can accept? We have seen that a central goal of CSR is to explain religion as such a phenomenon. According to Dennett, to view religion as a natural phenomenon is to view it as involving structures, processes, organisms, and events that obey the laws of physics and chemistry. Furthermore, it is because religion is natural in this way that we can explain it

(Dennett, 2006, 24-25).[12] This seems plausible. How else could religion be explainable than in virtue of being part of the natural world and therefore within the scope of an empirical investigation? It does not seem necessary to commit to physicalism to appreciate this.

If one instead includes in the notion of religion that it is a supernatural phenomenon, then given methodological naturalism, religion is no longer something we can explain. But that is not something a theist must necessarily do when pondering the possibility of an explanatory science of religion. Since the naturalness of religion (in the sense Dennett specifies) does not logically rule out theism, it seems possible for theists instead to embed this naturalness thesis and an open attitude to explanation in a larger theistic framework. From such a perspective, the role of CSR and related research could be to give us information about the nuts and bolts of a divine plan (see for example Barrett, 2004, cf. Visala, 2011). These nuts and bolts relate to structures and processes empirical science can investigate, while the divine plan is beyond the scope of such an investigation and must be worked out philosophically. Breakage of the first spell then does not necessarily lead to breakage of the second. That is, religion need not be "ruined" in virtue of being explained. If a theistic perspective on CSR can be plausibly sustained, then theists can even welcome the notion that science can explain religion, thus revealing proximate insights about a greater divine plan. Can religion thus re-emerge undamaged or even bolstered from having been scrutinized by explanatorily greedy scientists? As I shall argue, there are still serious challenges for attempts to demonstrate this.

What should we make of the tensions between CSR and theism that we have found? Those who argue in favour of a conflict-model must show that in the light of the tensions illustrated above (and further examples of such tensions) and despite the presence of serious reconciliation-attempts (illustrated above), it is in principle not possible to reconcile CSR and theism. To defend conflict (in this sense) is to defend a very strong claim. A more moderate and plausible thesis is that reconciliation is contentious and requires work but is in principle achievable. By contrast, those who argue in favour of reconciliation need to deal with the above-discussed and further challenges. They must, as a start, show that full-blooded CSR and theism can be reconciled in the sense that we can coherently think of both as true. But this seems a bit meagre for a reconciliation-advocate. This advocate might therefore also want to go further and show how CSR scholarship can be *harmonized* not only with mere theism but with a broader range of theistic theses and beliefs, thus bringing CSR together with a richer theistic framework. Then, there is a need to introduce further elements of what I call "theist religion" and to show

[12] This is not the only way to think of religion as natural. For a survey of further ways to talk about religion as natural, see De Cruz and De Smedt (2015, 30-32).

how CSR not only is logically compatible with but *hangs well together* with these. I shall argue *against* such attempts and thus also against the notion that CSR and theist religion are mutually agreeable (see chapter-outline below for details). I aim to show that their relationship instead is a disharmonious one. The normative implications in senses (2) and (3), which I shall seek to establish, show this. The next two sections introduce current debates about implications in those two senses. First, a remark about implications in sense (1). For the purposes of the present investigation, we let it remain an open question whether CSR and theism are or at least can become logically compatible and therefore in principle are reconcilable. We will not, in other words, try to resolve the issue of whether the normative implications in sense (1) that we have seen thus far indicate an irresolvable conflict. The primary goal is to look at normative implications in senses (2) and (3). We will see that implications in those two senses obtain even if we assume that the conflicts so far discussed are resolvable. CSR and theist religion have a disharmonious relationship even if we assume that CSR and theism are logically compatible.

4. What are the implications for the plausibility of theistic arguments?

That the findings of science can have impact on specific arguments for the existence of God is well-known. For example, research in physics is relevant in an appraisal of the fine-tuning argument, and research in evolutionary biology is relevant for William Paley's teleological argument. But what relevance does current *cognitive science* have for theistic arguments or theses? This is a major topic in the present investigation. Cognitive research on religion has normative implications for what I call theist religion in sense (2) if the claims and insights it provides are relevant to consider when assessing the plausibility of philosophical arguments or theses connected to the proposition that God exists. The present section provides examples. Richard Swinburne (2004) argues that the existence of consciousness, our ability to become aware of moral truths, facts about the structure of the universe, and religious experiences provide evidence for the existence of God. This is part of a larger cumulative argument for the conclusion that God probably exists. Alvin Plantinga (2000) defends the view that *if* God exists, then he has probably endowed us with a cognitive faculty that enables us to have non-inferential and immediate knowledge of God. Consider also a recent attempt by Justin Barrett and Kelly James Clark (2010, 2011) to show that findings in CSR provide empirical support for the notion that we have something like this faculty. This is an attempt not only to accommodate current

cognitive science to a theistic framework but to demonstrate how the former supports the latter. If insights from CSR bear on an appraisal of these positions, then it has the normative implications we are looking for. We shall see that this is the case with respect to that part of Swinburne's argument which concerns morality, and with respect to the thesis that we have a "god-faculty". Arguments *against* the existence of God are also relevant. They are connected to theism given that they attack that proposition. I will therefore also treat relevance of CSR for them as establishing normative implications in sense (2). The problem of evil is a well-known challenge that theism faces. The problem is how to reconcile the existence of a perfectly good, omnipotent, and omniscient deity with the existence of evil. There are also arguments from evil to the conclusion that God does not exist. They take evil as a premise and conclude that theism is false (or highly unlikely) (Taliaferro, 2004, 299). We shall consider what relevance CSR has for these arguments. Does such research tell us something of importance for the problem of evil? According to John Teehan, it does. Teehan (2016a) argues that findings in CSR (broadly construed as incorporating research on evolved morality) about what he calls the "cognitive science of evil" create what he calls a "cognitive problem of evil". In virtue of creating this problem, such research challenges the viability of attempts to integrate findings of CSR into theistic frameworks. This is therefore highly relevant. Furthermore, it seems that the cognitive problem of evil can add force to arguments from evil against theism. We here have further examples of possible normative implications in sense (2) for theist religion.[13]

The implications CSR has for theist religion in sense (2) relates strongly to what relevance such research has in the *philosophy of religion*, where appraisal of arguments or theses tied to theism is a central activity. This relevance is distinct from a logical bearing on the truth value of theism. If input from CSR can contribute in an appraisal of philosophical arguments for or against theism, then such research has distinct implications for theist religion. These implications deserve focus in their own right. It is not obvious that an appeal to empirical cognitive science helps us appraise theistic philosophical arguments and theses, given that these often focus on issues that are outside the scope of empirical science altogether or involve issues unrelated to cognition. Consider for example conceptual analysis of the concept of God, discussions of the nature of the Trinity or the existence of an afterlife, the issue of whether religious beliefs must be rationally defensible in order to be justifiably held, whether God's omnipotence is compatible with human free will, or the argument that the universe is fine-tuned

[13] Findings that bolster an argument against theism also have implications in sense (1), given that they bear negatively on the likely truth of theism. However, my primary focus will be on the distinct sort of relevance that pertains to philosophical arguments and theses.

for giving rise to complex life. These and similar topics, which philosophy of religion focuses on, are highly abstract and need not touch on results of the empirical study of religion.

Moreover, scholars in each of these two areas might also resist attempts to bring them together. Both philosophers and empirically oriented scholars of religion might want to highlight what they take to be the mutual strong independence of their respective areas. Viewing these fields as entirely unrelated seems to keep their autonomy intact and to allow a neat division of intellectual labour where empirical scholars study *belief in God* (among other things) and philosophers ask *whether theism is true* (among other things). This is something the former group will highlight to indicate that what they do is not theology or philosophy but science. Philosophers might also highlight this distinction to uphold or defend the autonomy of their field, which could seem threatened by the rapid developments in the empirical sciences. The differences between them seem obvious: while the former conduct studies and test claims, the latter engage in an abstract enterprise with the use of more or less pure reasoning from the armchair. Why mess up this neat picture by bringing empirical science into the philosophy of religion?

The broadly naturalist trend in philosophy of the last few decades has now also begun influencing the philosophy of religion. Importantly, introducing naturalism challenges the notion that this area is completely distinct from empirical studies of religion. What does naturalism mean here? Philosophical naturalism is a normative view about how philosophy should proceed.[14] According to this view, philosophy is not a purely a priori approach logically prior to empirical science. Philosophers should tentatively accept important scientific discoveries and start philosophizing in a landscape that the best current science describes.[15] Naturalists in philosophy thus use science as a *resource*. One feature that nevertheless makes philosophy distinct from science for the naturalist is that it asks *normative* questions, not only descriptive ones (see Godfrey-Smith, 2003, 149-155).

Given that naturalists see empirical research as relevant in philosophy, we should ask in what sense empirical results are relevant in a naturalistic philosophy of religion. Helen De Cruz and Ryan Nichols (2016, 3-4) describe two ways in which empirical research can be philosophically relevant, which both provide paths naturalistic philosophers of religion can take. *Experimental philosophers* conduct empirical (largely survey-based) studies to test their

[14] This means it is not the same as "strict naturalism" or methodological naturalism.

[15] Philosophical naturalism thus to some extent includes a thesis about what there is, not in the sense that philosophy provides any of its own or any final answers. Philosophical naturalism involves a tentative view of reality based on science, which is open to revision in light adjustments in existing theories.

claims. *Empirical philosophers* base their work (at least in part) on results of empirical research (that someone else carried out). Applied to religion this provides us with *experimental philosophy of religion* and *empirical philosophy of religion*. An example of the former is a study Helen De Cruz and Johan De Smedt carried out in order to look at how religious belief affects how one rates the strength of arguments for or against theism (2016, 119-142). An example of the latter is the comprehensive review of theistic arguments De Cruz and De Smedt (2015) provide, where they bring research in cognitive science to bear. They consider for example arguments from the appearance of design in nature, religious experiences, supposed miracles, and the phenomenon of human morality (here touching on Swinburne's moral argument). The approach in this review is to establish what kinds of intuitions that guide these arguments and then to investigate what CSR-research tells us about these intuitions. This provides a methodology in which CSR-research is relevant in the philosophy of religion because of how it sheds light on intuitions in philosophical argumentation (see De Cruz and De Smedt, 2015, xiii-xvii, 4-5). De Cruz and De Smedt call this approach "moderate naturalism". It includes methodological naturalism and the thesis that the justification of beliefs depends on the reliability of the process by which they arise.[16] This naturalist approach exemplifies empirical philosophy of religion insofar empirical results are relevant in philosophy.

De Cruz and De Smedt do not argue that CSR has strong impact on the arguments they review but rather that it has extensive relevance by shedding light on underlying intuitions in a broad range of cases. For those interested in carrying out a normative analysis of the soundness of arguments, much more than an investigation of CSR is therefore needed. How then, if at all, is CSR a substantive resource for the philosopher of religion? De Cruz and De Smedt (2015) describe such research primarily as an inspiration for natural theology rather than as a substantive resource by which we decisively can settle issues. They do, nevertheless, provide an example of how an appeal to CSR can warrant a strong conclusion about a theistic argument. This example concerns an argument from widespread belief in God to the existence of God as a best explanation. CSR challenges this argument by providing an alternative naturalistic explanation of why belief in God is widespread (De Cruz and De Smedt, 2015, 184-185, see also Visala, 2011, 160 for this assessment of this kind of argument).

De Cruz and De Smedt provide many examples of how a study of human intuitive thought is relevant in a discussion of theistic arguments, but the impact that cognitive science has on such arguments is in my view greater than what they argue. Moreover, it is negative:

[16] This thesis is known as *reliabilism.* Chapters 6 and 7 deal with this view.

findings in cognitive science *undermine* arguments from morality and the thesis that we have a god-faculty, while to some extent *strengthening* the argument from evil *against* theism. In the final chapter, we return to the phenomenon of widespread belief in God. We shall here see how this phenomenon arguably provides a basis for criticism of atheism. I argue that a naturalistic explanation of widespread belief in God provides an auxiliary tool that aids the advocate for atheism by providing a response to this criticism. This provides yet another example of how CSR can have negative impact on theism, but this time more indirectly by coming to the aid of the atheist advocate rather than by having direct impact on a specific argument.

Notice that if cognitive science has strong implications for theistic arguments, then that by extension also indicates possible implications for the status of *theistic beliefs*. If for example we obtain reasons to dismiss a theistic argument, then that has implications for the epistemic status of beliefs based on it. Then, implications in sense (2) also by extension have possible implications in sense (3), the topic to which we now turn.

5. What are the implications for the epistemic status of theistic beliefs?

That CSR has the sort of relevance for beliefs that I hinted to in closing the previous section is not obvious of course. For some, it is not only unobvious but categorically mistaken to even suggest such a thing. According to *fideism*, religious utterances and beliefs are not appropriate targets of a critical rational scrutiny. To believe in God is to have a certain basic orientation or commitment which does not involve the formation of beliefs that require rational normative support. Instead, one's orientation relies solely on faith, a sort of non-reasoned trust whereby one has taken a leap into the unknown.[17] If that view is correct, then even if CSR has strong impact on arguments that support (or undermine) theism, such research will nevertheless not have epistemic relevance for theistic beliefs. This is because such beliefs (if that is how we should talk about the religious orientation mentioned above) involve faith and therefore require no external support from arguments. We will not deal with this view. Instead, we assume that

[17] Plantinga (1983, 87) distinguishes two forms of fideism. According to what he calls "moderate fideism", "faith" does not depend on reason for normative support. The moderate view does not dismiss the use of reason to investigate or defend faith but sees it as superfluous. According to "extreme fideism", reliance on reason in religion is in principle wrong. Reasoning is not only superfluous or alien to faith but the act of relying on it is a hindrance to faith. This is because by trying to justify one's theistic beliefs using reason one fails to trust God, and such trust relates closely to faith. Extreme fideism therefore implies that there is a conflict between faith and reason.

there are such things as religious beliefs and that they can and should be critically scrutinized and assessed with respect to epistemic justification. This is a central assumption in contemporary analytical epistemology of religion, an enterprise with which we will engage. Granting this assumption, we immediately encounter questions about what sort of factors that have a bearing on the normative status of beliefs in general. As will become clear below, what theory of such factors one supports has potentially strong implications for how much epistemic relevance we plausibly can attribute to CSR.

I indicated in the previous section that CSR possibly sheds light on some factors that bear on the epistemic status of beliefs, namely reasons, arguments, and theses tied to theism. In this case, the epistemic relevance for beliefs comes by an indirect route that goes via implications in sense (2). However, this is not the only way for CSR to have such epistemic relevance. As I will argue, reasons, arguments, and theses on which one can base beliefs are not the only factors that bear on their epistemic status. What other factors play such a role and what does CSR tell us about them? This is a major topic in the present investigation.

The present section provides an overview of the main positions in the current debate concerning what epistemic relevance CSR has for theistic beliefs. The section focuses on *three distinct views* that have emerged recently: debunking, neutrality, and positive relevance (see also De Cruz and Nichols, 2016, who provide their own overview). We will see that they involve very different views on what kind of factors that bear on the epistemic status of beliefs and that they then apply those views to theistic beliefs. In section 6, I indicate which view I argue for and how I intend to show that CSR has epistemic relevance given this view.

Let us start with the view that recently has gained the most attention, debunking. Guy Kahane (2011, 104) describes an evolutionary debunking argument (EDA) as an argument that appeals to an evolutionary story about the origins of a group of beliefs in an attempt to undermine the justification of those beliefs. Kahane's discussion focuses on the usage of such arguments in metaethics but there are also examples of such arguments in the philosophy of religion. Such arguments try to discredit beliefs by pointing to facts about how they arise, either ultimately in terms of the origin of faculties that produce them or proximately in terms of how they arise in the individual. If CSR-scholarship motivates an EDA that targets theistic beliefs, then such research has normative implications in sense (3). The strength of these implications depends on the soundness of the argument and, at least when it comes to its relevance for our enquiry, the extent to which results of *CSR* rather than general evolutionary considerations do the debunking.

I indicated that debunking (targeted at religion) has gained much attention recently. James Jones (2016, 1, 6-7) says that currently a "torrent" of attempts to use CSR-research to debunk religion faces us. Visala (2011, 170) describes a "plethora" of debunking-arguments based on CSR. This does not mean, surprisingly, that a large number of critics of religion employ such arguments. Though there are exceptions, it turns out that critics of debunking who defend theism from such possible attacks based on CSR are the primary purveyors of these EDAs. Consider for example Thurow (2013) Leech and Visala (2011), Clark and Barrett (2011), Barrett (2007), Murray (2010), and Clark and Rabinowitz (2018), who all present and criticize self-constructed debunking arguments. An exception to this pattern includes Jong and Visala (2014), who respond to a debunking argument Wilkins and Griffiths (2012) defend.[18]

Here is J. Barrett's construal of the debunking line:

> "Selection pressures (operating on either groups or individuals) have led to various dispositions and propensities in human minds that happen to give rise to religious belief. Religious beliefs are, therefore, accidents or byproducts of evolution. As such, religious beliefs cannot be trusted. Belief in gods amount to cognitive illusion" (Barrett, 2007, 62).

This argument starts with the claim that religious beliefs are evolutionary by-products and concludes that they therefore are examples of *cognitive illusion*. This is an example of an EDA. An evolutionary story of the origin of religious beliefs allegedly shows that such beliefs amount to a cognitive illusion. This conclusion is stronger than what normal EDAs conclude, which is that beliefs *lack justification*. The conclusion seems to entail the falsity of theism, if we interpret "cognitive illusion" as belief in something illusory. J. Barrett's objection to this argument is that to claim that gods are illusory is to step outside of science (which is involved in the premise of the argument) and enter metaphysics (Barrett, 2007, 63). He thus diagnoses a possible conflict between science and religion by pointing to a transgression against methodological naturalism whereby one reads an atheistic (metaphysical) thesis into the by-product thesis.

Another and more straightforward response is that the argument seems fallacious. Unless we add the premise that beliefs that are evolutionary by-products lack justification, the conclusion does not follow. It is far from obvious what the basis for such a premise is. It also seems fallacious to argue that a belief is false because of how it arises. This is an example of

[18] But this pattern seems to be changing. For example, Goodnick (2016) and Braddock (2016) have defended EDAs. Van Eyghen (2018) responds to these. The present investigation does not discuss these contributions to the debate.

the so-called genetic fallacy. The *truth* of a belief does not depend on how it arises but whether the proposition the belief involves is true. In the case of theist belief, only the existence of God (and therefore not any fact about human psychology) can make that proposition true.[19]

John S. Wilkins and Paul E. Griffiths (2012) present and endorse an EDA. They argue in favour of the notion that the evolution of belief-forming faculties can lead to scepticism with respect to many but not all the beliefs they produce. Such scepticism affects religious and evaluative beliefs but not common-sense beliefs. The idea is not that evolutionary considerations give us reasons to believe that religious and evaluative beliefs are false but that they can undermine our reasons to believe they are true. Why does evolution not similarly affect common sense beliefs? Common sense beliefs are about middle-sized mundane objects we encounter in our surroundings. The reason evolution does not undermine them is that they arise from processes that *track truth*. An appeal to evolution can show this. We can establish a probable link between variation in ability to form true common sense beliefs and variation in fitness. Wilkins and Griffiths argue that those who get things right in the domain of common sense have a better chance of survival and reproduction than those who do not. Thus, there is a "bridge" from getting things right in this domain and reproductive success (Wilkins and Griffiths, 2012, 134-135). However, there is no evidence of selection for ability to track *religious truth*. No link between getting things right with respect to religion and reproductive success can be established. This undermines the epistemic status of religious beliefs (Wilkins and Griffiths, 2012, 142-143).

Wilkins and Griffiths do not appeal to specific claims or proposals in CSR when presenting their argument. Their discussion of such research is rather general. In the argument they present it is largely a story about truth tracking and the lack of evidence for such tracking in the formation of religious and evaluative beliefs that does the debunking. It is therefore unclear how strongly CSR specifically rather than general evolutionary considerations discredits religious beliefs. Furthermore, it is not clear that such general considerations warrant a strong normative conclusion concerning religious beliefs. For this to be plausible, one must assume a defensible normative principle according to which lack of truth-tracking undermines (or is incompatible with) justification. If not, the argument seems to face the same problem as the one J. Barrett's argument faces, namely a missing premise.

We will see that failure to track truth is plausibly an epistemically relevant issue and that such failure in a process that produces religious beliefs casts doubt on such beliefs. I shall

[19] This is so if we assume a *realist view of truth*. For an example, see Alston (1996). Moore and Scott (2007) provide an introduction to the realism/anti-realism debate in the philosophy of religion.

argue that the focus should be on whether the process is *reliable*, a notion related to truth-tracking.[20] Furthermore, I argue in favour of a normative theory according to which reliability is necessary for justification. Here we find a normative principle from which a negative epistemic conclusion follows from lack of reliability. But this is not yet sufficient for a debunking argument based on *CSR*. It also seems necessary that we bring in something more specific. A plausible debunking-argument requires empirical detail about the processes by which theistic beliefs arise. The present investigation applies insights found in CSR to provide such detail and discusses what these insights can tell us about the reliability of the responsible processes. This shows how CSR has strong implications in sense (3) for theist religion.

Let us now turn to another central position in the literature, which rejects debunking and argues in favour of a very different thesis. Some of those who reject the debunking view argue that CSR-research is epistemically irrelevant with respect to theistic beliefs. They seem to endorse the following thesis:

> CSR has no epistemic relevance for theistic beliefs.

I call this the "epistemic neutrality thesis" (ENT). Let us consider two different arguments for ENT. This allows us to see the importance that choice of epistemological theory has. The distinguishing factor is whether the epistemological theory the argument involves is *internalist* or *externalist.* I start with the former. According to internalism, epistemic justification is *entirely* dependent on factors internal to the subject. This means factors that are cognitively accessible or that are part of the mental life of the subject (Conee and Feldman, 2011, 54-55). Internalists view justifying factors as things we can think about and base beliefs on, such as reasons, evidence, arguments, and experiences. Such factors are internal. Internalism seems involved in the following statement about how to show that theistic beliefs are unjustified:

> "That conclusion can only be obtained by first assessing the reasons people give. Furthermore, let's suppose that a thorough assessment of those reasons shows that they do not support religious beliefs. Well, then, religious beliefs will have been shown to be unjustified and the findings of CSR about how people actually form their religious beliefs will be superfluous. So, the findings of CSR don't by themselves show religious beliefs to be unjustified and once we have assessed the reasons that people actually give for their religious beliefs, the findings of CSR wont' affect that assessment in any way." (Thurow, 2013, 93).

[20] For details see the chapter-outline below.

Here an assumption is that the *only* appropriate method for evaluating beliefs involve assessment of the reasons people have. Notice also the rather severe view that religious believers who lack good reasons to believe are unjustified.[21] Consider also a remark Jonathan Jong and Aku Visala make in a criticism of debunking-arguments:

> "As we have seen, if S fails to provide arguments or evidence for p, her belief that p is unjustified anyway (unless her belief that p counts as properly basic, if there are any such beliefs); the questionable provenance of S's belief that p isn't doing the work here. If, however, S *can* provide arguments or evidence for her belief that p, then it is justified, regardless of the psychological origins of the belief." (Jong and Visala, 2014, 249).

Here we again see the view that theists who lack good reasons (or evidence and arguments) to believe in God lack justification (with the qualification that such beliefs are not properly basic, a notion I will deal with extensively later). Both statements say that epistemic evaluation of beliefs should involve a discussion of the reasons or evidence subjects have rather than what cognitive science tells us about how their beliefs arise. In other words, facts about how beliefs arise do not bear on their justification. Only reasons, evidence, and arguments do.

Here is *an internalist argument for ENT* based on these and similar views:

I-ENT

P1: Only factors that are internal[22] to the subject have epistemic relevance.

P2: CSR does not shed light on internal factors (like reasons, evidence, or arguments).

C: CSR has no epistemic relevance for theistic beliefs (ENT).

For internalists this is a straightforward argument in favour of irrelevance. We will see that one major problem it faces is that P1 is problematic. Externalism is a more plausible position. If we give up P1, then the step from P2 to C is unwarranted. P2 is also problematic. As indicated above, CSR seems relevant in assessment of at least some arguments for the existence of God. In chapter 4, section 5, I argue that it is relevant with respect to moral arguments. Since these

[21] Externalists are more permissive: they allow that external factors such as the reliability of belief-forming processes can justify the beliefs of subjects that lack access to good reasons.

[22] "Internal" here means internal the way required for epistemic relevance in accordance with internalism. See chapter 4, section 3 for further discussion.

and other arguments operate like reasons on which theists can base their beliefs, then CSR can shed light on internal factors.[23] Given that both premises are problematic, I-ENT is problematic.

A move to externalist epistemology opens up for a wider range of epistemically relevant factors and thus seems to increase the chances that CSR has epistemic relevance. According to externalism, *not all* epistemically relevant factors are internal (Goldman and McGrath, 2015, 42). But despite this permissiveness with respect to epistemically relevant factors, it is not obviously true that CSR has epistemic relevance given externalism. We now look at an externalist argument *against* the contention that CSR has such relevance. This argument concedes a central epistemically relevant factor recognized by many externalists, namely the process by which beliefs arise, i.e. the proximate etiology of beliefs. Given this, scientific insights concerning belief-production is clearly of epistemic interest. Isn't CSR then of at least potential relevance in assessing the epistemic status of theistic beliefs? How then do we obtain an externalist argument for ENT?

Consider the question, "why do you believe in God?" Suppose we take this to be a request for an explanation. Should we turn to CSR for an answer? According to Aku Visala (2011, 127-128), we should not. CSR does not tell us how theistic beliefs arise in individuals. If this is correct, then CSR seems not to have epistemic relevance with respect to theistic beliefs, even given externalism. This motivates an *externalist argument for ENT*:

E-ENT Preliminary

P1. How theistic beliefs arise is an epistemically relevant factor.[24]

P2. CSR cannot tell us how theistic beliefs arise.

C. CSR does not have epistemic relevance in virtue of telling us how theistic beliefs arise.

C does not entail ENT because externalists also (like internalists) see internal factors like reasons, evidence, and arguments as epistemically relevant. They only claim that those are not the only epistemically relevant factors. C is therefore consistent with CSR having epistemic relevance for beliefs in virtue of shedding light on internal factors. We therefore use P2 from I-ENT as the fourth premise in the following, stronger externalist argument for ENT:

E-ENT

[23] However, it is not clear that theists normally base their beliefs about God on specific arguments. I argue in chapter 3 that the processes by which theistic beliefs normally arise are more intuitive and less reasoned than that.
[24] Notice that Thurow, Jong, and Visala deny this premise when saying that the psychological origin of theistic beliefs is irrelevant.

P1: How theistic beliefs arise is an epistemically relevant factor.

P2: Internal factors like reasons, evidence, and arguments are epistemically relevant.

P3: CSR cannot tell us how theistic beliefs arise.

P4: CSR does not shed light on internal factors (like reasons, evidence, or arguments).

C: CSR has no epistemic relevance for theistic beliefs (ENT).

P4 in this argument is the same as P2 in I-ENT and therefore faces the same challenge, the possibility that CSR is relevant with respect to at least some arguments on which theists can base their beliefs. The central premise in E-ENT is P3. Let us look closer at it. It is an ambiguous statement. We can take it to mean that CSR cannot tell us how theistic beliefs arise in *specific individuals* or that it cannot tell us *by what processes theistic beliefs arise* in general. We can furthermore distinguish a strong and weak reading of these two claims. On a strong reading, "cannot" indicates a principled impossibility, and on a weak reading, it indicates that something is presently not the case. I start by focusing on strong readings. A question arises concerning what "cannot tell us" plausibly should mean in this context. This concerns explanation. CSR allegedly cannot *explain* how theistic beliefs arise. But should P3 be construed as stating that CSR cannot *provide a sufficient explanation* of how theistic beliefs arise or that such research cannot *contribute to a sufficient explanation*? To make P3 sufficiently robust to offer support for ENT, we should construe it as denying that CSR can contribute to explaining individual beliefs.[25] The reason we should construe P3 in that way is that defending this denial seems necessary to offer significant support for ENT. This is because it is not necessary for the epistemic relevance of CSR with respect to individual theistic beliefs that it can fully explain such beliefs on its own. It is, as we shall see below, sufficient that it can contribute to doing so. That is something those who favour E-ENT therefore should not accept.

[25] It follows that CSR also fails to provide a sufficient explanation of such beliefs. This is because if X does not contribute to explaining Y, then X does not sufficiently explain Y. But the converse is not true: If X fails to sufficiently explain Y, it could still contribute to sufficiently explaining Y. The processes CSR reveals seem to contribute in this way and therefore roughly resemble what John Mackie (1974, 62) calls an "INUS" condition: "an *insufficient* but *non-redundant* part of an *unnecessary* but *sufficient* condition" {emphasis original}. These processes are insufficient because mere cognitive processes without social and cultural input do not produce specific, theistic beliefs. They are non-redundant because they nevertheless contribute in significant ways to the production of these beliefs in many cases. But they are unnecessary, since it is possible for theistic beliefs to arise through other processes. It is for example possible to form theistic beliefs using reflective and reasoned means. For this reason, it is in principle possible also to explain the formation of theistic beliefs by asking subjects for their reasoned grounds to believe. However, in light of CSR, this does not apply to the majority of theistic beliefs. Instead, most theistic beliefs arise in ways subjects do *not* access by introspection, and they are often wrong about the factors that cause their beliefs (for this argument, see Ch. 3, section 2). Finally, the processes CSR reveals are parts of sufficient conditions for the production of theistic beliefs: general CSR-processes in combination with specific inputs from the social and cultural environment cause such beliefs (see Ch. 3, section 3, 4, and 5 for more).

P3 faces challenges given this construal, since we have grounds to reject the thesis that CSR cannot contribute to explaining individual beliefs. CSR can plausibly contribute to explaining individual belief-forming episodes by providing knowledge of general processes by which religious beliefs form. By combining such general knowledge with information about social and cultural peculiarities particular to individual cases, information we obtain by making relevant observations, we can obtain the basis for a sufficient explanation of these cases. In short, the combination of general and specific information allows us to subsume individual cases as tokens of general process. An analysis of the epistemic status of the general process will then have implications for the epistemic status of its individual tokens. For this reason, CSR has potential epistemic relevance, given externalism, if it *contributes* in this sort of way.

Let us also consider the issue of whether CSR can explain by what general processes theistic beliefs arise. Because of the role of general knowledge in explaining individual cases and because of the epistemic implications such explanations can have, we should construe P3 as also denying that CSR sufficiently can provide such general knowledge. As just mentioned, general knowledge about belief-formation contributes to explaining individual cases and this opens up for potential epistemic significance. Therefore, unless the externalist advocate for ENT denies that CSR in principle can provide general knowledge, a troubling path to epistemically relevant explanatory contributions is opened up, thus undermining ENT. But P3 faces problems on this construal too. Not only does CSR research (or at least those parts of it I will focus on) aim at providing general explanations of how a wide range of religious beliefs form, as we shall see (especially in chapters 2 and 3), work in the field has already achieved important results in this regard by exploring how human cognition interacts with input from social and cultural surroundings. This at least indicates that in principle such a general explanatory goal is within the scope of CSR.

Thus far then we have grounds to construe P3 in E-ENT as denying (in principle) that CSR can contribute to explaining individual beliefs and that CSR can sufficiently explain belief-formation in general. A successful, principled objection according to which, for example, it would be conceptually confused, strongly misguided, or categorically mistaken to apply cognitive science to explain beliefs seems to permanently block CSR-arguments against ENT. This therefore seems the most important goal for those who support E-ENT. However, given the challenges such an objection faces, the advocate for E-ENT might be tempted to make the weaker objection that CSR at present merely fails to help explain how individual theistic beliefs form, or that it fails to fully explain by what general processes this happens. This brings us to the weak readings of P3. One could grant in principle that CSR can provide the sort of insights

an externalist would see as epistemically relevant, whilst denying that any such insights presently emerge from the field. This move does not permanently block externalist arguments against ENT but, if successful, it supports that thesis tentatively. We should therefore construe P3 as also denying that CSR presently contributes to explaining individual episodes and that it presently explains general belief-forming processes. The present investigation argues against also this, on the assumption (specified above) that central contributions to CSR are true. The above-mentioned results, assuming they emerge from a field whose central contributions are in fact true (or approximately true or likely true) indicate that CSR presently helps explain individual beliefs and offers explanations of the general processes by which they arise (for a detailed account, see Ch. 3, sections 3, 4, and 5) – or at least likely does this. This means that also the weaker readings of P3 are wrong. In summary, we can interpret P3 as making one or more of the following 4 claims:

1. It is in principle impossible for CSR to contribute to explain how theistic beliefs form in each individual.
2. CSR does not presently contribute to explain how theistic beliefs form in each individual.
3. It is in principle impossible for CSR to explain by what processes theistic beliefs form.
4. CSR does not presently explain by what processes theistic beliefs form.

Some brief examples will illustrate that scholars make claims that relate, roughly, to these components of P3. For example, Aku Visala argues that the sort of explanation of religion we find in CSR, which focuses on universal and general trends, does not provide tools that allow us to reveal how specific individuals form their beliefs. Such explanations are "useless" as tools to reveal why a given individual has a specific religious belief (Visala, 2011, 127-128). Pascal Boyer warns against using CSR to explain individual belief-forming episodes, contending that social science does not allow us to identify the unique causal events that lead to belief in individual cases (Boyer, 2002, 366-367). These statements motivate a view which commits at least to statement 2, possibly also statement 1.[26] The idea seems to be that appealing to CSR in an attempt to explain individual belief-forming episodes is misguided and conceptually confused, since CSR does not help explain this.

[26] Visala has (personal communication) indicated that Visala (2011) does not commit to the claim that it is impossible, in principle, for CSR to explain how individual beliefs form. Some of the things Visala (2011) but also Boyer (2002) say about CSR and explanation (illustrated above, and to be discussed further in chapter 3) nevertheless motivate the view that it is misguided and wrongheaded to apply CSR in an endeavour to explain individual beliefs.

According to Helen De Cruz and Johan De Smedt, research in CSR does not plausibly motivate EDAs because it does not allow us to specify the responsible processes by which theistic beliefs arise. One important reason is that any attempt to specify processes correctly faces the *generality-problem*, the problem of specifying process-types at their right level of generality (De Cruz and De Smedt, 2015, 193-194). These statements motivate a view which commits at least to statement 4, and possibly also, at least if the generality-problem is deemed unsolvable, statement 3.

I will respond to these two views (for details, see chapter-outline below), arguing that if we want to explain how theistic beliefs arise in individuals it is not necessary to provide detailed knowledge of the unique causal events that occur in each one of them. It suffices that we specify the relevant process-types and then correctly subsume individual episodes as their tokens. To subsume correctly, we must obtain information which is specific for the given episode we want to explain, and then show how this episode is a token of the general process. Since CSR provides the general knowledge which is one crucial component in this sort of explanation, statements 1 and 2 are wrong: CSR can in principle contribute and does, if true, at presently contribute successfully to explain individual episodes. Concerning statement 3, I argue that CSR can in principle provide a sufficient explanation of a set of process-types by which theistic beliefs form. It does this by providing insights into general processes and mechanisms by which the mind processes information provided by the social and cultural setting. Moreover, assuming central contributions to this field are true, statement 4 is also false: CSR presently explains by what processes theistic beliefs form, or to qualify a bit, it presently offers insights into at least some such processes. This provides sufficient grounds to deny P3 in E-ENT.

We have seen how an internalist and an externalist argument challenge the notion that CSR has epistemic relevance for theistic beliefs. The former narrows the range of epistemically relevant factors to a set that seems to be outside the scope of CSR-scholarship. The latter assumes a wider range of factors but points to an explanatory inadequacy in such scholarship to establish epistemic irrelevance. In light of the problems these two arguments face, we have reasons to reject them. CSR therefore seems at least potentially epistemically relevant for theistic beliefs. As mentioned earlier, I will argue that this research also has actual epistemic relevance for theistic beliefs (see the chapter-outline below for further details).

Before we get to that, I want to flag a further specific view in the literature. This view, related to Alvin Plantinga's epistemology (and mentioned in section 4), denies ENT by viewing CSR as having epistemic relevance for theistic beliefs yet disagrees with debunking by contending that this relevance is *positive.* Kelly James Clark and Justin Barrett (2010, 2011)

bring CSR-research and a position known as *reformed epistemology* (RE) together and argue that the former provides empirical support for central contentions in the latter. According to RE, belief in God is basic, meaning that it arises spontaneously and automatically from the normal functioning of a cognitive faculty. According to Clark and Barrett, CSR supports the contention that many unreasoned religious beliefs arise in a similar way. Such beliefs are "instinctive" in that they often arise automatically and without inference from the normal functioning of human minds (Clark and Barrett, 2010, 174). Clark and Barrett contend that CSR thus provides support for the existence of a sort of *god-faculty*. By "god-faculty", they mean that human minds working normally often produce beliefs in supernatural intentional beings (Clark and Barrett, 2011, 652). This leads to a positive relevance thesis because of how RE treats the epistemic status of intuitive beliefs that arise spontaneously and without inference:

> "If the position of Reformed epistemologists is correct that the automatic, non-reflective deliverances of our cognitive faculties should be treated as innocent until proven guilty, and if humans are naturally endowed with a god-faculty or sensus divinitatis, then belief in the divine should be treated as innocent until proven guilty" (Clark and Barrett, 2010, 176).

CSR is showing us that theistic belief, or at least something close to this, i.e. belief in the "divine", arises in a way that RE deems sufficient for warrant or justification in the absence of independent reasons or evidence to support them. Thus, the position of Clark and Barrett seems to grant P1 and deny P3 in E-ENT. That is, it suggests that how theistic beliefs arise is an epistemically relevant factor and that CSR (by providing evidence for a god-faculty) tells us how such beliefs arise. Assuming RE is correct and that CSR actually tells us this, the relevance thus seems positive. The chapter-outline below points out how I will respond to this view.

We have now reviewed the main positions concerning epistemic implications for beliefs, what I take to be normative implications in sense (3) for theist religion. The next section provides an outline of the remaining chapters of the present investigation.

6. Plan for the remainder of the investigation

We can group the remaining chapters together in three pairs and one concluding chapter. Chapters 2 and 3 constitute the first pair. The main topic in this part of the investigation is what CSR tells us about how theistic beliefs arise. Chapter 2 serves as an introduction to research in

CSR, which presents what I find to be the most compelling work. I also criticize some alternatives and respond to some recent criticism of this work. We shall focus on theses defended by Stewart Guthrie, Justin Barrett, Pascal Boyer, Will Gervais, Joseph Henrich, and Lee Kirkpatrick. A central idea we will discuss is that religion is an evolutionary by-product of how various cognitive systems solve fitness-relevant tasks. We shall focus on an agency-detector, a system devoted to handling social exchange, adaptations for obtaining socially transmitted information, and an attachment-system. The main purpose of the chapter is to shed light on our general human susceptibility to believe in supernatural intentional agents. Chapter 3 explores the formation of theistic beliefs. I make a distinction between how *people become theists* and how *theistic beliefs* form. I argue that research in this area sheds important light on these topics in general and in extension also helps account for individual cases. To be a theist is to be prone to form beliefs about God in a range of situations. I call this propensity the *theistic disposition*. This disposition arises gradually as the result of how a culturally distinct concept becomes psychologically relevant. To account for this, our general susceptibility to believe in supernatural agents and an account of how our minds use information in distinct, theistic cultural contexts are relevant. The theistic disposition is thus a profound result of a set of general processes coming together in a specific theistic cultural context. The combination of the same processes will in other contexts lead to propensities to form beliefs about different deities, or to entities like ghosts, fairies, or ancestor-spirits. This account challenges P3 in E-ENT, thus undermining one central source of support for the epistemic neutrality thesis (ENT).

How can CSR-scholarship be relevant in the philosophy of religion? Chapters 4 and 5 are devoted to this issue. Chapter 4 considers theistic evidentialism (TE), which I treat as the thesis that only evidence can justify the belief in God, and that this belief in fact is justified by good evidence. The internalist view of justification this position involves is problematic. Cognitive processes we cannot introspect and facts about our environment are plausibly also epistemically relevant factors. This conclusion (which is not based on CSR but general epistemological considerations) challenges TE and I-ENT (the internalist argument to neutrality). The second part of TE is the thesis that there is good evidence for the existence of God. One source of purported evidence comes from moral arguments for the existence of God. CSR has a bearing on such arguments, including one Richard Swinburne defends. A problematic feature of these arguments is their unwarranted scepticism with respect to the ability of naturalistic science to account for morality. Morality is also a relevant topic in an argument John Teehan (2016a) presents, which in my view reveals that a cognitive understanding of morality leads to serious explanatory challenges for those theistic theses that

attempt to integrate current cognitive science into a larger religious framework. The discussion of these issues reveals that CSR has normative implications in sense (2) for theist religion.

Chapter 5 considers reformed epistemology (RE) and focuses on the claim that we have a god-faculty. RE is plausible, in my view, in regarding facts about the functioning of cognitive faculties as epistemically relevant. However, the claim that we have a god-faculty is problematic. I shall reveal how an imagined advocate of the god-faculty who by stipulation finds CSR convincing and who wants to retain central theses in RE faces a dilemma. On the one hand, this advocate can choose an interpretation based on Plantinga (2000), according to which the god-faculty is specialized to produce beliefs about God. The problem is that CSR is strongly at odds with this thesis. On the other hand, the advocate can instead apply an interpretation based on Clark and Barrett (2011), arguing that the god-faculty is a general tendency to form beliefs in the supernatural. The problem here that this god-faculty cannot deliver what RE promises, namely properly basic belief in *God*. The discussion of this dilemma reveals normative implications in sense (2): CSR bears negatively on RE, a thesis connected to theism.

What relevance does CSR have in evaluation of the epistemic status of theistic beliefs? Chapters 6 and 7 deal with this question. Chapter 6 defends an externalist virtue-epistemology called agent-reliabilism. Agent-reliabilism is the thesis that knowledge consists of true belief arising from an epistemic virtue. Furthermore, justification requires that an epistemic virtue has been relevantly involved (Greco, 2007, 177, 2010, 10-12, 71). Epistemic virtues are reliable belief-forming dispositions: they tend to lead to true and to avoid false belief with respect to a relevant range of propositions when the subject is in a range of appropriate circumstances and conditions. Such dispositions have their basis in natural cognitive faculties or in thinking-habits the subject gradually acquires (Greco, 2007, 211-217, Greco, 2010, 3-16, 71-90). In chapter 7, we investigate whether the theistic disposition, which is an example of a thinking-habit that subjects gradually acquire, meets necessary requirements for being an epistemic virtue. I argue that the theistic disposition is a culturally distinct token of a broader type of disposition that produces beliefs in cognitively attractive agents in general. Among the circumstances and conditions relevant for this type we have reasons to see multifarious cultural and social conditions. When we implement this type in a broad range of such conditions, the result is the production of different religious beliefs. Strong grounds that also theists will accept indicate that most of the purported gods, spirits, and demons involved in those beliefs do not exist. But then, even if theism is true, this disposition does not tend to lead to true and to avoid false beliefs the way virtue requires. This has negative implications for the theistic beliefs that arise in

theistic cultural contexts. Our investigation here reveals how CSR has normative implications in sense (3) for theist religion by having relevance in an evaluation of the epistemic status of theistic beliefs.

Chapter 8, the concluding chapter, takes stock of the normative implications the investigation has revealed and makes a point about an additional role CSR can play in criticism of religion.

Chapter 2. The Human Susceptibility to Concepts of Supernatural Agents

> "As Richard Dawkins once remarked, there seems to be one simple process whereby people 'get' their religion and that is *heredity*. After all, the best way to predict people's religion is to find out their parents' religion. Now this does not mean, obviously, that being a Buddhist or a Mormon comes with a particular chromosomal configuration. Dawkins' tongue-in-cheek remark was meant to emphasize something religious people often take for granted but is a source of constant amazement to outsiders: that people generally adhere to the specific religious commitments of their community and ignore other variants as largely irrelevant." (Boyer, 2002, 365).

1. Introduction and chapter-outline

Where do concepts of supernatural agents like deities, spirits, fairies, demons, and ghosts come from and why are they so widespread in human cultures? How do human minds mentally represent such concepts? Furthermore, why do some of them occur in religious belief-systems while others are widely entertained as mere fictional ideas? These and more questions will occupy us in the present chapter. For answers, we turn to the Cognitive Science of Religion (CSR), by which I mean a loosely delineated area in the study of religion whereby scholars attempt to provide an explanation of religion as a natural phenomenon in terms of human

cognition and evolution. The chapter serves as an introduction to CSR, which expands on the brief account I gave in the previous chapter. The presentation revolves around a selection of representative, early contributions and some recent developments relevant to these. We shall also deal with some controversies that have arisen in the field. Our goal is not to provide a broad introduction which aims to cover as much as possible of this rapidly developing field but to explore what some central contributions to such scholarship tell us about our *general susceptibility to belief in the supernatural*. This lays a foundation for our discussion in chapter 3 about how subjects in theistic cultural contexts gradually become prone to form beliefs about a culturally specific deity, God. Here is an outline of the present chapter.

As we saw in chapter 1, section 1, we shall explore theories according to which religious beliefs and behaviors are by-products of the normal functioning of evolved cognitive systems. These theories commit to the view that religion is an *evolutionary by-product*. Even though still dominant, this view has increasingly been criticized by those favoring *adaptationism.* Section 2 presents this controversy and provides grounds for favoring the by-product view. While recognizing the complexities of this issue and the need for continued investigation, we shall then for the purposes of the present investigation assume that religion is a by-product.

Our next topic concerns the human tendency to see human features in the natural world. To survive, our ancestors needed to be highly sensitive to the possible presence of human agents in their surroundings. Given the high costs of failure to detect such agents, our minds became prone to infer their presence in ambiguous situations. This is the central theme in section 3, which presents the work of Stewart Guthrie, a foundational scholar in CSR who focuses on our tendency to attribute human features to the non-human world. In the dynamics of this process we find the basis for the unintended human creation of agent-concepts, some of which eventually became culturally successful and highly elaborated notions like Yahweh, the Feathered Serpent, Zeus, and Allah. We here have a possible answer to the question of where such ideas ultimately come from. Importantly, these are concepts of *humanlike* agents with whom we can communicate. We next look at Justin Barrett's development of Guthrie's thesis. Barrett defends what has come to be known as the HADD-thesis, the view that we have a hyper-active or hyper-sensitive agency-detection device ("HADD"). We look at what sort of device this is and what possible role it can play in religious belief-formation.[27]

In section 4, we turn to the work of Pascal Boyer on the structure of the concepts we find in religion. These concepts partly consist of socially transmitted ideas subjects obtain

[27] We return to the topic of agency-detection in chapter 3, section 5, where we look at how "explanatory theistic beliefs" can arise from this kind of process. We again discuss it in chapter 7, section 2.

during social interactions and partly also of innate intuitive thought by which we carve reality into various parts. Boyer argues that religious concepts involve small violations of such intuitive thought. Section 5 presents a theory defended by Boyer and Justin Barrett (among others) about why such slightly counterintuitive agent-concepts are widespread in human cultures. This theory, known as MCI theory, says that concepts with this structure are widespread because we are susceptible to such concepts, and that this susceptibility in turn reflects their cognitive attractiveness. I shall defend this theory from recent criticism offered by Benjamin Purzycki, Aiyana Willard, and Stewart Guthrie. At this point we have answers to the questions of why some types of concepts rather than others occur widely in human cultures. We also want to know why some of them become embedded in religious belief-structures while others remain mere cultural fictions. Section 6 introduces an important factor that helps explain an obvious and easily observable yet puzzling aspect of religion: people tend to believe in the deities of *their culture* rather than those of other cultures. This is puzzling because the latter are equally attractive. The solution, which Will Gervais and Joseph Henrich purport to provide, concerns biases and heuristics involved in how subjects obtain socially transmitted information and use it as a basis when forming their own beliefs. I shall look at a tendency to conform to norms, behaviors, and customs central to the group. These processes reflect the functioning of adaptations to culture, the humanly constructed environment of socially transmitted information.

Section 7 discusses the work of Lee Kirkpatrick on attachment, an adaptation for enhancing proximity between offspring and caregivers. Kirkpatrick argues that God[28] can become an *attachment-figure* and that when this happens, the psychology of attachment strongly influences how the subject thinks about and relates to God. Attachment-theory of religion provides a bridge that connects the topics of cultural evolution and proximate psychology. It is therefore highly important given our aim to shed light on how people become religious. Based on the contributions to CSR that we look at in the present chapter, we have answers to the above-specified questions and a foundation for asking how people become theists and how theistic beliefs arise.

[28] Not necessarily what I mean by "God" (See Ch. 1, section 2) but related to it, given the many references to Christian beliefs.

2. A prelude: The by-product/adaptationism debate

Many of the early, foundational contributions to CSR share the view that religion is an evolutionary by-product, and this view continues to be defended by central scholars in the field today. Consider for example Boyer (1994), Lawson and McCauley (1990), Guthrie (1995), Atran (2002), Barrett (2004), Kirkpatrick (2005), Pyysiäinen and Hauser (2009), Bloom (2010), Baumard and Boyer (2013), and Boyer (2018). But there are also a range of alternatives to the by-product view. For example, Wilson (2002a), Johnson and Bering (2010), Johnson (2016) and Bulbulia and Sosis (2011) defend the thesis that religion is an *adaptation*. Moreover Norenzayan (2013) and Norenzayan et al. (2016) combine cultural adaptationism with the by-product view in an attempt to explain both the emergence of large-scale societies and of institutionalized religions featuring "Big Gods". These alternative hypotheses challenge the hegemony of the by-product view in the field, and they play a central role in current discussions of the nature and ultimate origin of religion. However, we shall not base our investigation on these alternatives but instead assume that the earlier, and still quite influential by-product view is true (recall, our investigation looks at implications of CSR assuming central theories in the field are true). The present section, which is a prelude to a more detailed presentation of a selection of theories, briefly considers some alternatives to that view. I argue that they face challenges in the light of which we lack decisive grounds, at least for the purposes of the present investigation, to depart from the by-product view and instead assume the truth of adaptationism. I end with a remark about what possible implications the choice of an adaptationist versus a by-product view has for a central argument later in the present investigation.

For the purposes of the present investigation, a given theory of religion is adaptationist if it maintains that religion is a *biological adaptation* and it is a by-product theory if religion is seen as *a by-product of biological adaptations*. If the theory combines the latter view with the view that religion is a cultural adaptation evolved through cultural group selection, then we shall refer to it as a *hybrid* view of religion.

Let us start by considering the *supernatural punishment theory* Jesse M. Bering and Dominic Johnson promote. Johnson and Bering (2010) and Johnson (2016) introduce the evolutionary puzzle of cooperation and its ubiquity in our species as their starting point. The massive amount of cooperation characterizing our species is ancient. It is not a phenomenon which occurs only in present day society but something which stretches far back into our past, prior to the existence of large-scale societies. It is widely recognized as an evolutionary puzzle (see e.g. Boyd and Silk, 2015, 434). One reason this phenomenon is puzzling is that it extends

beyond kin and small groups of individuals with whom one forms reciprocal relationships. This means the mechanisms of *kin-selection* and *reciprocal altruism* are not on their own sufficient to account for it (Boyd and Silk, 2015, 428-429, and see chapter 4, section 5 for further discussion). Johnson and Bering argue that these and further mechanisms, such as *indirect reciprocity*, in which cooperating with a non-related individual with whom one does not have a reciprocal relationship is fitness-conducive since it helps establish a good social reputation, do not explain the extent of human cooperation. This is because its scope includes cooperation even under conditions where one's identity is hidden from both recipients and possible by-standers (Johnson and Bering, 2010, 27-28). Johnson and Bering argue that religion is an adaptation for cooperation which accounts for how such behaviours nevertheless evolved. Religion thus solves the puzzle of cooperation. Religion is on this view something very concrete: a "precise, cognitive mechanism" which induces cooperativeness by way of creating a fear of supernatural punishment for failing to cooperate (Johnson and Bering, 2010, 29). This mechanism evolved by natural selection as a response to new selection-pressures which were introduced when complex language and Theory of Mind (the ability to interpret behaviour by attributing mental states; more on this term below) emerged in the hominid mind. The reason selection-pressures changed when they were introduced to the cognitive makeup of our ancestors is that it allowed subjects to damage the reputation of non-cooperative individuals by spreading negative social information (i.e. "gossip") about them, thus lowering their fitness. For this reason, a cruder selfishness which previously in our evolution was fitness-conducive now became much costlier (Johnson, 2016, 153).

To illustrate how the religious adaptation could have evolved, Johnson and Bering (2010) and Johnson (2016) ask us to consider three possible behavioural strategies and propose a comparative analysis of their relative adaptiveness in this new and more socially complex environment. "Ancestral individuals", without language or Theory of Mind, simply act out selfish desires without the use of the sort of social subterfuge more cognitive sophistication allows. In this environment, "Machiavellian individuals", who are also prone to behave selfishly but in various novel and psychologically sophisticated ways, would outcompete the former individuals because of their ability to exploit social information in devious ways. However, the Machiavellians would at the same time be liable to incur the cost of punishment when detected. By contrast, "God-fearing individuals" also have selfish urges, but their fear of supernatural punishment for selfishness helps keep them in check, thereby inducing more cooperative behavior. The cost of this strategy is missed chances to exploit others. This is according to Johnson and Bering a minor cost compared with those one incurs when behaving

selfishly. To be more specific, the "God-fearing" individuals have an advantage to the Machiavellians, since the cost of missed opportunities to cheat is on average lower than the total cost of punishment for cheating (Johnson and Bering, 2010, Johnson, 2016, 165-168). This tells us that fearing supernatural punishment is adaptive. As Johnson so memorably puts it,

> "Gods offer a way to reform the selfish ape inside us so that we can avoid the wrath of hairless and brainy apes around us" (Johnson, 2016, 157).

What are "gods"? [29] The deity of the Abrahamic religions (who I simply refer to as "God") is one example, but it is not the only one. Fear of supernatural punishment can be delivered also by witches, sorcerers, ancestors, and a range of deities of varying stature and power (Johnson and Bering, 2010, 31).

Given that humans have an adaptation for cooperation which works by producing beliefs in supernatural punishment for selfishness, we can expect such beliefs not only to be widespread today but also quite ancient. For example, they should precede the emergence of large-scale societies with their institutionalized, moralizing, religions (more on this below), where the prime deities clearly are thought of as punishing moral deviousness. Indeed, as Johnson points out, given the time-scale in which natural selection works, we need to look at how things were as far back as 200,000 years ago (2016, 51). Did our ancestors at that time believe in supernatural punishment for non-cooperative behaviour? The study of present-day indigenous societies provides a window to this past, and indicates tentatively that they did, argues Johnson (2016, 51-57). But this view is being challenged. As for example an archaeologist working at Göbekli Tepe, an archaeological site created at the onset of the Holocene ("just" 11,000 years ago) argues, it is not clear whether the notion of a god existed at this time (Schmidt, 2010). It then seems even less clear whether such a notion existed 100,000, not to mention 200,000 years ago. Moreover, as Nicolas Baumard and Pascal Boyer (2013) argue in a discussion of beliefs of a much more *recent* origin than that, the gods or spirits of small-scale groups prior to the emergence of large-scale societies but after the end of the Pleistocene were thought of as *uninterested* in cooperation and morality. These entities were instead believed to be much more

[29] Given how often Johnson (2016) uses the term "God" in his discussion, one might wonder how this relates to how I use that term. Our usage is related, since Johnson relates "God" to the Abrahamic religions and points out that this, for example, is *not* a deity Hindus or Buddhists believe in (Johnson, 2016, 3-4). However, I employ classical theism, which attributes a specific set of attributes to God. This is not something Johnson does. Moreover, on Johnson's theory, the central belief which is said to produce fitness-benefits is about *supernatural punishment*, and Abrahamic god-conceptions are not the only ones which achieve that. For these reasons, we must clearly keep "God" the way Johnson uses this term and "God" the way I use it apart.

interested in whether subjects carried out the right rituals and showed the correct amount of devotion. Moreover, this is true even of the greater deities of some successful recent large-scale societies, such as Sumer, Rome, or the Aztec empire (Baumard and Boyer, 2013). In light of this, it would be puzzling if supernatural entities of much more ancient origin would have been moralizing and concerned with deterring selfishness.

Another challenge is that it is not sufficiently clear that our ancestors (for example 200,000 years ago) were as prone to crude selfishness as the above account indicates. Was a religious adaptation necessary as a mechanism to ensure that groups managed to cooperate? Is the only plausible alternative to the "God-fearing" subjects devious "Machiavellians" who presumably have no regard for their conspecifics and who would do anything to exploit them? Not if we follow the account of morality Michael Tomasello (2016) proposes. According to this account, the challenges to cooperation increased cognitive sophistication produced were solved by the cultural evolution of social norms and a variety of psychological mechanisms for making subjects abide by them (Tomasello, 2016, 85-128). None of these work by way of creating beliefs in the supernatural. For further examples of alternative accounts where religion at best plays a peripheral role, consider Joyce (2006), Boyd and Silk, (2015, 417-435), Laland (2017), Boyer (2018), Buchanan and Powell (2018), and see chapter 4, section 5 in the present investigation for further discussion. These accounts allow us to propose further evolved behavioural strategies. For example, based on Tomasello's account, we can imagine a norm-abiding type of individual who is not only motivated to monitor others and punish their non-cooperativeness, but who believes he has an obligation to be loyal to his culturally delineated group and who has the ability to curb his own selfish desires out of this sense of obligation. This type of individual is able to judge his behaviour morally and will feel guilt or shame for failure to abide by the relevant norms (Tomasello, 2016, especially pp. 107-111). It is not clear that the "God-fearing" individuals would outcompete this sort of individual, who clearly is not akin to the devious Machiavellians, but who instead is quite prone to cooperate with in-group conspecifics. In light of these considerations, we have grounds to be sceptical of the current example of the supernatural punishment thesis.

Let us now turn to a hybrid theory of religion defended by Norenzayan (2013) and Norenzayan et al. (2016), which also defends the centrality of fear of supernatural punishment.[30] The main difference from the above theory is that while Johnson and Bering propose that such fear is a *biological adaptation*, Norenzayan sees this as a culturally attractive idea which

[30] For yet another example of a supernatural punishment theory, see Watts et al. (2015).

prevails in groups as the result of how it enhances their cohesiveness and stability (Norenzayan, 2013, 136). Another difference is that while Johnson and Behring are interested in the ultimate origin of religion in biological evolution, presumably sometime back into the Pleistocene, Norenzayan (2013) and Norenzayan et al. (2016) are interested in much more recent cultural developments occurring at the onset of the Holocene. Their theory assumes a by-product view of religion (for this, see Norenzayan, 2013, 8) and emphasizes the need for a cultural evolutionary process to account for these developments. The following two puzzles provide a starting point for Norenzayan et al. The first is the emergence of complex large-scale societies during the Holocene and the second is the emergence of the notion of big, powerful, moralizing, punishing, gods, henceforth, "Big Gods", during the same epoch (Norenzayan et al. 2016, 2-3). The proposed solution involves the claim that pro-social religion, that is, religion involving Big Gods, has evolved culturally through *cultural group selection*. In short, religious beliefs and practices revolving around Big Gods helped increase the level of cooperation and trust within some groups. This increase in cooperation and trust in turn contributed to the success of these groups in competition with other groups. The success of these groups then led to the cultural success of the notion of Big Gods (Norenzayan, 2013, Norenzayan et al., 2016).

How central were beliefs in Big Gods in contributing to the emergence of large-scale societies? Norenzayan et al. point out, correctly in my view, that this sort of religion is not necessary for enabling large-scale cooperation. Moreover, they continue, neither is pro-social religion sufficient for creating enough cooperation to allow large-scale societies to arise. Instead, it is one of more "likely causes" for the emergence of prosocial, large-scale societies (Norenzayan et al. 2016, 3). Norenzayan et al. allow for the possibility of many different causal factors being at play, and even allow that in some cases, large scale cooperation emerged without belief in Big Gods being involved at all (Norenzayan, et al., 2016, 9). Something this leaves us curious about is the extent to which belief in Big Gods compared with other factors has contributed to the emergence of large-scale societies in general. If Big Gods are not crucial to establish such a result, and a range of other factors also have been at play, there is the danger that the latter replaces the need for the former as an explanation or that the contribution of Big Gods is an anomaly to a general pattern in which entirely different factors do the explaining.

The two main lines of evidence for the Big Gods thesis is experimental evidence indicating that religious primes (reminders of the notion of Big Gods) induce pro-sociality (this evidence will be considered in chapter 4, section 5) and archeological evidence indicating that beliefs in moralizing deities have contributed to the emergence of ancient large-scale societies. Above I argued we should be skeptical of the notion that belief in moralizing supernatural

punishers were widespread prior to the Holocene. But is there any evidence for belief in Big Gods during the transition to the Holocene or subsequent to this transition? Norenzayan (2013) points to the archeological site of Göbekli Tepe in present day Turkey as possible evidence for such a transition. This site, which is believed to have been involved in religious rituals, was created more than 11 000 years ago, prior to the emergence of cities, by what were probably hunters and gatherers (Norenzayan, 2013, 118-120). Norenzayan argues that this site indicates the possibility that a propensity to worship Big Gods, "motivated people to take up early forms of farming, and not the other way around" (Norenzayan, 2013, 120). This is significant because the emergence of organized farming, that is, agriculture, is a central event that marks the transition into a new epoch (the Holocene). But what did the subjects at Göblekli Tepe believe? As mentioned above, the archeologist Claus Smidt, who has done extensive research in this area, says that we currently do not know what they believed, and it is unclear whether the concept of a god even existed during early Holocene (Schmidt, 2010). But Schmidt does not rule out belief in gods completely. The central elements at the site are large stone pillars which according to Schmidt depict humanlike entities, and we cannot, in his view, rule out that they depict gods of some kind (Schmidt, 2010, 253-254). But given this uncertainty about what the subjects at the site believed, it seems uncertain whether belief in Big Gods – a specific subset of deities consisting of powerful, punishing, moralizing deities, was involved here. Let us therefore instead turn to some more recent large-scale societies. When it comes to some of the greatest ancient civilizations, such as those of Mesopotamia, the Aztec and Mayan empires, and ancient Greece or Rome, Nicholas Baumard and Pascal Boyer argue that subjects here did *not* believe in Big Gods. The gods of these societies where uninterested in moral transgressions but cared primarily about being served by their subjects in the right ways through rituals and sacrifices (Baumard and Boyer, 2013, for a response, see Norenzayan et al. 2016, and for further discussion, see Teehan, 2016b). It seems then that subjects managed to cooperate sufficiently to create a broad range of sophisticated societies without the help from belief in Big Gods. That this is a general pattern in the emergence of large-scale societies is further supported by research conducted by Harvey Whitehouse and colleagues. Whitehouse et al. (2019) looked at more than 400 societies that have existed during the last 10,000 years and investigated the statistical relationship between social complexity and the extent to which the god or gods are moralizing and punishing. The earliest evidence for such a god they found was in Egypt, 2800 BC, a relatively recent moment in the Holocene. By contrast, the Big Gods hypothesis leads us to expect such gods much earlier. Moreover, Whitehouse et al. also found, again contra the predictions the Big God hypothesis leads to, that belief in such deities *do not precede* but *tend*

to follow the emergence of large-scale societies (Whitehouse et al. 2019). This does not, as Whitehouse et al. (2019) point out, show that Big Gods played no role at all in contributing to societal stability. However, it undermines the notion that Big Gods have played a central causal role in bringing about large-scale societies. A final problem is this: suppose, contra what the current evidence suggests, that the emergence of large-scale societies in fact were strongly associated with the emergence of belief in Big Gods. That does not tell us that the latter contributed to *causing* the former. It is also possible (as Norenzayan et al. 2016, 9-10 recognize) that widespread belief in Big Gods arose as a side-effect of the emergence of large-scale societies (for a by-product account of the emergence of moralizing religion, see Baumard et al. 2015, and Boyer, 2018). In light of these considerations, we have grounds to be skeptical towards the Big Gods hypothesis, at least if we by this mean an ambitious attempt to offer an explanation of the formation of large-scale societies in general.

We have made a brief foray into some recent work on religion, and I have offered some critical remarks, which by no means are intended as providing conclusive grounds to dismiss the above theories. Our goal is to introduce research in CSR and to indicate some challenges that representative alternatives to the earlier but still dominant by-product view face. Let us now turn to that view. We start with a concrete example which neatly illustrates how it differs from the above theories.

Pascal Boyer (2002) argues that we do not have cooperation and morality because we have religion. Instead, we have first evolved cooperation and morality, and this nudges us spontaneously to evaluate the behaviors of others morally. This tendency makes it easier for concepts of deities, produced as by-products of our minds, to latch on to our moral intuitions, thus providing us with the idea that gods care about morality. As Boyer puts it, we do not have morality because we have religion. We have religion because we already have (evolved) morality (Boyer, 2002, 193). The content of religious moral ideas therefore reflects the content of an (evolved) morality which does not hinge on religion. In support for this, intuitive moral beliefs have been found to be the same whether one holds religious beliefs or not (Baumard and Boyer, 2013, Pyysiäinen and Hauser, 2009, and see Ch. 3, section 3, and Ch. 4, section 5).

Now to the general picture: A range of cognitively attractive concepts have gradually emerged in human natural history as the result of the functioning of evolved cognitive systems which solve adaptive challenges unrelated to religion. Once such concepts are around, exposure to them leads subjects to form religious beliefs and to engage in various religious behaviors (Boyer, 2003, Barrett, 2004). Different aspects or components of what one might call religion are thus by-products of the functioning of a variegated collection of systems with distinct

evolutionary origins (Boyer, 2003). For example, this includes an agency-detector (Guthrie, 1995, Barrett, 2000, 2004), Theory of Mind (e.g., Barrett, 2004, Boyer, 2002, Norenzayan, 2013, 17-18, Kirkpatrick, 2005), a cheater-detector (Boyer, 2002), systems for forming intuitions about physical and mental entities (Bloom, 2004), an attachment-system (Kirkpatrick, 2005), a "precaution system" (Boyer and Lienard, 2006), and mechanisms for detecting kin (Kirkpatrick, 2005, 249-251).

Consider for example Theory of Mind (henceforth "TOM"), an evolved system or set of systems for interpreting behaviors in terms of mental states such as beliefs and desires (Boyer and Barrett, 2005, 205). Scholars like Boyer (2002), Barrett (2004), Kirkpatrick (2005), McCauley (2013), and many others see this system as centrally involved in religious beliefs and behaviors, and studies support its involvement (Willard and Norenzayan, 2013). Yet, it is likely to be much more ancient than religion, and its adaptive function is not to provide religious beliefs. According to Andrew Whiten, TOM can be related to a larger set of evolved systems which together make up a "deep social mind", consisting also of a capacity for language, cooperation, and culture, which has evolved in our species during the Pleistocene as a solution to adaptive challenges related to the need to interact with and handle complex tasks together with conspecifics (Whiten, 2013, cf. Barrett, 2015, 128-152). If this system nevertheless gives rise to religious beliefs, then these are by-products of its functioning. Another example is mechanisms for detecting and interacting with kin. Their adaptive function is to enhance inclusive fitness by helping individuals detect and act altruistically towards kin. We share this trait with other primates, and it is probably evolutionary quite ancient (Boyd and Silk, 2015, 173-206). Yet, kinship-related thought can come to expression in religious contexts as thoughts about ancestors or gods or about co-religious believers as "fictive kin", such as when referring to a deity as "father" or co-religious believers as "brothers and sisters" (Kirkpatrick, 2005, 246-251). Such thoughts are by-products of kin-detection. Similar accounts can be provided for the other mechanisms mentioned above.

The main point is that the mechanisms found to give rise to various aspects of religion, and which are involved presently in religious beliefs and behaviors, evolved as solutions to problems that had nothing to do with religion. Lee Kirkpatrick captures the by-product view nicely, stating that,

> "Religious beliefs are constructed, shaped, and maintained by a host of psychological mechanisms and systems – including the attachment system – that all evolved much earlier in

> the (pre)history of our species for more mundane purposes, but that have been "co-opted" in more recent human history in the service of religion" (Kirkpatrick, 2005, 214).

We shall now explore some of these systems and how religious ideas trigger them. In this way we shed light on the general human susceptibility to a broad range of beliefs in the supernatural. This presentation, which continues our defense of the by-product view, prepares the ground for the account in chapter 3 of how subjects in theistic cultural contexts become theists.

But before we get to that, let us briefly ask whether assuming an adaptationist alternative to the by-product view would impact on central arguments in the present investigation. One issue is this. A central premise I make use of in chapter 7 in building my case against the epistemic justification of theistic beliefs is that subjects in different cultural contexts are likely to believe in different supernatural agents. One might wonder whether this premise hinges on the truth of the by-product view or whether one could defend it also from an adaptationist perspective. I believe one could. The adaptationist view clearly recognizes the manifest cultural variability religion exhibits and purports to explain why it obtains. Here are two examples to illustrate. In a defense of adaptationist theories of religion from by-product critics, Richard Sosis argues that it is not the case that if something is an adaptation, then it operates in the same way in all contexts. Rather, genetic predispositions we all share need not be expressed in a particular environment or in the same way in different environments (Sosis, 2009, 326). So, religion can be an adaptation yet vary in how it is expressed in varying environments. Moreover, we saw above how Johnson and Bering (2010) emphasize that fear of supernatural punishment can come to expression in the form of belief in a broad range of supernatural entities. Johnson points out that the variety of postulated agents by which such punishment is thought to be delivered is enormous (Johnson, 2016, 59), and he argues further that cross-cultural variety in beliefs, rather than homogeneity, is exactly what an adaptationist theory of religion leads us to expect (pp. 92-93). In light of this, the premise I make use of in chapter 7 concerning the massive variability of religious beliefs seems possible to defend also from an adaptationist perspective.

3. From Stewart Guthrie's anthropomorphism thesis to Justin Barrett's "HADD"

The anthropologist Stewart Guthrie is one of the earliest contributors to a cognitive science of religion. Guthrie (1980) applies a cognitive perspective to religion which was further developed in *Faces in the Clouds* (1995). Guthrie belongs to an "intellectualist" tradition that views religion as an attempt to understand, influence, and explain the world (1995, 10, 21). This attempt involves intuitive inferential strategies in perception whereby we attribute agency to aspects of reality that lack agency. Guthrie grounds religion in a general proclivity in the hominid mind to employ inferences that *anthropomorphize*, that is, attribute human features to non-human aspects of the natural world (Guthrie, 1995, 62-64, 188). Guthrie's central claim is that religion is anthropomorphism. Given that this is a universal psychological phenomenon rather than something specific to religion, an adequate explanation of religion requires that we look broadly at the functioning of the human mind (Guthrie, 2017, 80).

The human tendency to *animate* and to *anthropomorphize* comes to the fore in perceptually ambiguous situations (Guthrie, 1995, 39-41). These are situations where we encounter things whose nature is unclear to us. That we anthropomorphize means that we do two things. We animate nature by thinking it is alive, and then we assume that what is alive is humanlike (Guthrie, 1995, 62). Animism can be as simple as the misidentification of an autumn leaf as an insect. Non-human animals also do it. Guthrie for example mentions chimpanzees that make aggressive displays as if challenging thunderstorms. He points out that it is not absurd to assume a storm is alive. A storm can move about, it has a sort of coherent structure or outline, makes scary noises, and leaves its path devastated (Guthrie, 1995, 52). These features could trigger an evolved agency-detecting capacity because it seems to suggest something like self-propelled movement and intentionality. Psychological studies suggest that humans have a proclivity to make animating inferences when encountering many non-living phenomena. For example, one study suggests that children intuitively interpret dots on a computer monitor that move in certain ways as agents (Atran, 2002, 61). According to Guthrie, what occurs in such cases as well as in encounters with various things in nature that seem alive is the outcome of a normal human tendency to interpret natural phenomena as exhibiting agency. An important premise in his argument for the centrality of this tendency is the view that perception is a top-down process involving interpretation and that when interpreting, we are prone to "bet" on the presence of the most significant phenomena, namely living agents (Guthrie, 1995, 52).

We can think of such bets as what the evolutionary psychologist H. C. Barrett labels "inductive bets", that is, guesses about a fitness-relevant feature of the environment. Various bet-making strategies are according to Barrett part of the functioning of the mind and they tend to pay off when the organism inhabits environments that sufficiently resemble those

environments in which the mind evolved (Barrett, 2015, 23-25). Statistically during hominid evolution, entities moving in a seemingly purposeful fashion would have been alive. A strategy which in a modern experimental setting might lead to mistaken inferences would therefore on average in the sort of environment in which we evolved tend to pay off.

The evolutionary process favors strategies that yield an average good payoff even if they occasionally lead to mistakes. Why should this be so? Guthrie points out that a lot of information relevant to the possible presence of human and nonhuman animals in our natural surroundings is ambiguous. This is because there is so much stealth and deceit in nature. Both humans and animals therefore make mistakes and think that something non-living is living, thereby animating the world. In other cases, the mistake is failure to detect something alive. Such failures are often the result of successful stealth and deceit (Guthrie, 1995, 50-51). The occasional mistakes allowed by the strategies that evolution favors are not random. Let us explore why.

There is an asymmetry between the average costs of falsely thinking something is present and falsely thinking that something is not present. The cost of the former mistake is that we worry about an agent that is not there. By contrast, failure to detect a real agent can in some cases be *fatal*. It is therefore, to avoid incurring this cost, reasonable for our minds to assume order and complexity (i.e. agency) when information is ambiguous (Guthrie, 1995, 45). The biologist Richard Dawkins can help us appreciate this dynamic. Dawkins (2016) calls this sort of asymmetry in the costs of making mistakes the "life/dinner" principle. To illustrate how this could play out in nature and how different the costs can be for different species, he describes an arms race between foxes and rabbits where selection pressures make the former increasingly better at hunting rabbits and the latter increasingly better at escaping predation. Dawkins points out that the cost of failure is different for the fox and the rabbit. The cost for the fox who fails to detect and capture a rabbit is a missed dinner, but the cost for the rabbit who fails to detect a fox is its life. By contrast, the costs of falsely inferring the presence of a rabbit or fox is expenditure of energy by chasing or escaping from something which is not there (Dawkins, 2016, 100).

Let us look closer at these two types of errors. When it comes to the presence or absence of an agent, a "false positive" (also called "type 1 error") is an inference that an agent is present when no agent is present. This inference assumes more order and complexity than what is present. A "false negative" (also called "type 2 error") is a failure to infer that a real agent actually is present. This mistake fails to attribute the correct amount of complexity present (Barrett, 2015). Thus, as Dawkins so succinctly puts it, false positives infer non-randomness

(i.e. order and complexity) when one encounters randomness, and false negatives infer randomness (lack of order and complexity) when one encounters non-randomness (Dawkins, 1998, 171-172).

Here are some further examples provided by Dawkins, which illustrate the logic at play here. A farmer who thinks a sacrifice to a deity brings rain makes a false positive. He assumes a causal relationship (i.e. non-randomness) between his sacrificial action and the weather. A bird that makes the bet that a twig is a stick caterpillar assumes more order than what is present, again a false positive. By contrast, the mistake the farmer makes if he fails to detect a real relationship between a weather-related deity and the weather is a false negative. Moreover, a bird that mistakes a caterpillar for a twig makes the same mistake, namely to fail to correctly infer agency (non-randomness). The cost of a false positive is unnecessary expenditure of time and energy (performing a pointless sacrifice or trying to capture a twig). But the cost of a false negative for the farmer might be high (divinely produced bad weather which destroys his crops) and for the bird negligible (a missed caterpillar) (Dawkins, 1998, 172-174).

The farmer's reasoning about a deity in this example is not an integral part of a general intuitive inference-strategy in the human mind. Such specific thoughts about a deity and its relationship to the weather are probably more particular to that case (and as we shall see later, they require exposure to a religious concept). By contrast, the intuitions we shall focus on currently are rather automatic (and general). But these examples neatly illustrate *the general logic* involved in such thought, which is to nudge us to think in ways that on average increase fitness more than alternative ways. Here the asymmetry in cost between type 1 and type 2 errors is crucial. A general principle for how this affects the evolution of cognitive strategies is the following. If the cost of false positives is *higher* than the cost of false negatives, then we can expect a betting-strategy with a high threshold for inferring the presence of the phenomenon. But if the cost of false positives is *lower* than the cost of false negatives, then we can expect a strategy with a low threshold for inferring the presence of the phenomenon. In these two cases, the strategies help maximize fitness by reducing the likelihood that one makes the costliest mistake (Dawkins, 1998, 176, cf. Haselton, Nettle, and Andrews, 2005).

In the case of hominid agency-detection, Guthrie (1995) argues in favor of a strategy whereby the threshold for placing a bet in the presence of the phenomenon is low. Our minds employ this strategy because the cost of false positives is in this case lower than that of false negatives. But Guthrie does not make strong claims about what specific system in the mind is responsible for creating this bias towards belief in the presence of agency.

The psychologist Justin Barrett (2000, 31, 2004, 32-33) interprets this bias as reflecting the functioning of what he calls a "hyperactive agency detection device" (HADD). He describes this as a system or mental tool for detecting agency (Barrett, 2004, 32-33). It is an example of a "categorizer", by which the mind determines the nature of what it has encountered (Barrett, 2004, 4-5). Guthrie interprets Barrett's HADD-idea as an attempt to cash out Guthrie's general idea about the human propensity to over-infer agency. This leads to the notion that when we encounter ambiguous phenomena, *HADD* triggers the inference that agents are present, thus helping us avoid the costly type 2 error (Guthrie, 2017, 76). What sort of system is this?

We have grounds to see HADD as an evolved functional system specialized to deal with the multifarious challenges that the presence of agents in our surroundings represents.[31] H. C. Barrett (2005) proposes, focusing in particular on the adaptive challenge that the presence of predators and prey represented in ancestral environments, that we possess a complex cognitive system evolved to process information concerning agents. This "agency-system" consists of a set of mechanisms for carrying out various tasks, such as detecting agents, discriminating between kinds of agents, and for attributing intentions to them (Barrett, 2005, 203-204). This system's agency-detector is according to H.C. Barrett highly sensitive to seemingly self-propelled motion and to cues that indicate whether an object reacts to events in the environment. A restricted set of what Barrett (2005) refers to as "behavioral signatures" trigger the agency-system to make further inferences about the assumed agent. Different signatures for example indicate whether it engages in pursuit or escape. We also possess a few evolved perceptual templates for specific kinds of animals such as snakes and spiders. Specific morphological and behavioral cues trigger inferences about their presence. H. C. Barrett finally points out that the agency-detector is especially sensitive to structures or cues that suggest the presence of eyes and the direction of eye-gaze. Eye-gaze indicates what the agent is looking at. This is highly fitness-relevant. For example, whether the agent has detected us or not is crucial information (Barrett, 2005, 204-209). The evolved system H. C. Barrett here describes provides a plausible basis for HADD, which J. Barrett (2004, 33) describes in similar ways as a functional system that responds to self-propelled movement and automatically produces intuitions about agency in fitness-conducive ways.[32]

[31] Not everyone in CSR will agree with this interpretation. For example, Mark Andersen rejects the thesis that agency-detection is carried out by HADD and argues that we instead should interpret this process as an instance of a domain-general perceptual process, "predictive coding", which involves Bayesian inference (Andersen, 2019).
[32] In a recent, critical discussion, Neil Van Leeuwen and Michiel Van Elk (2018) argue that HADD theory faces a problem: current neuroscientific evidence does not point to *one specific agency-detecting module* but to various neural systems devoted to agency-related tasks. This evidence is in their view incompatible with the thesis that HADD is a module (Van Leeuwen and Van Elk, 2018, 5-7). But firstly, it is not clear that J. Barrett and other

Thus far then, HADD has a plausible basis. When it comes to the functioning of this system, however, we have grounds to make some possible adjustments to the picture J. Barrett presents. This concerns the notion that our agency-detector is "hyper-active" or "hypersensitive". The evidence H. C. Barrett reviews does not point to an agency-detector which makes us *excessively* prone to false positives but suggests that we are relatively reliable in detecting agents (Barrett, 2005). H. C. Barrett (2005, 213, cf. Maij. et al. 2019) stresses that agency-detection instead involves a *moderate* bias in the direction of over-attribution of agency. By being somewhat "jumpy", we avoid costly false negatives. Nevertheless, there are also important limits to how readily our mind will infer the presence of agents on uncertain evidence, given the cost of "unnecessary vigilance" (Barrett, 2005, 205). This view relates closely to the general idea in evolutionary psychology that when a moderately biased solution on average is less costly than a more biased or a more accurate solution, then we can expect a moderately biased solution to be favored (Haselton et al., 2005, 726). Since we view the agency-system H. C. Barrett describes as the basis for HADD, it seems that HADD is prone to false positives (rather than negatives) but not in an excessive or indiscriminate fashion. In light of this, some possible adjustments to HADD-theory which take account of this might be in order. The reason I describe them as *possible* is that it is not entirely clear whether advocates of this theory see HADD as excessively prone to attribute agency. J. Barrett for example says that the term "HADD" is used primarily to highlight the fact that our agency-detector *sometimes* becomes hyper-active (Barrett, 2004, 31-32). That does not mean he contends that it always is.[33]

HADD-advocates see HADD in that way. For example, as mentioned, Barrett (2004) describes HADD as a "mental tool" for detecting agency. This tool allows us to identify whether or not something we encountered exhibits agency (Barrett, 2004, 4, cf. Barrett, 2005, 204). But Barrett (2004) and Barrett and Lanman (2008), who further develop HADD-theory, do not state that HADD is a single neuronal module with a specific localization in the brain. Second, evidence for numerous distinct systems governing various tasks related to agency is not incompatible with the notion that HADD is a module, since it is possible for modules to comprise submodules (Hirschfeld and Gelman, 1994, cf. Bechtel and Wright, 2009, and Gottschling, 2009, on the many different ways in which a module can be specified). Therefore, it is possible that HADD is functionally specialized to carry out the tasks Barrett (2004) and Barrett (2005) describe *and* that it comprises a set of more specific modules devoted to distinct aspects of these tasks. As for example H. C. Barrett argues, the agency-detector is made up of more specific mechanisms which respond to distinct inputs statistically related to the presence of agency (Barrett, 2005, 203-204). This underlying complexity does not mean the agency-detector is not a module in the sense of a functionally specialized system. It means that its overall functionally is served by more specific, functional systems. In light of this and given the range of evidence Barrett (2005) amasses, the alleged lack of evidence for a neuronally distinct, single system for agency-detection does not in my view seriously undermine HADD-theory.

[33] Barrett (2004) indicates that HADD becomes increasingly prone to false positives in ambiguous or scary situations. But do such situations actually accentuate this propensity? To test this, Maij et al. (2019) carried out an experiment whereby they exposed their subjects to inputs that induce a feeling of threat, and they then looked at whether this increased the propensity to falsely detect agents. They found that subjects to *some extent* tend to falsely detect agents when presented with ambiguous stimuli. However, inducing threat using various manipulations (including virtual reality) did not produce an increased bias towards false positives. In a discussion of the results, Maij. et al. (2019, 45-46) argue that the (moderate) bias towards falsely detecting agents is roughly but not entirely in line with HADD-theory. We are to some extent biased towards falsely detecting agents, but the

Let us next appreciate that Guthrie (1995) argues that we have a proclivity not only to animate but to *anthropomorphize*. In support of this, he points out that human agents are the most complex and significant phenomena we can encounter and that applying a human template to ambiguous input therefore yields the greatest payoff. Humans (and our pre-human ancestors) are for example capable of endless varieties of behaviors, some of which involve stealth and subterfuge and lead to the violent death of conspecifics. Given how ambiguous and dangerous the human presence therefore is for other humans (statistically in ancestral environments), readily applying a "humanlike model" makes evolutionary sense (Guthrie, 1995, 39-90, cf. Duntley, 2005). How readily do our minds apply that model? To be more specific, given that HADD provides the basis for the proclivity Guthrie describes, how prone to anthropomorphize is HADD? Given how sensitive this agency-detector is to very specific cues that suggest the presence of predators, prey, and specific animals (spiders and snakes), it does not seem to use a human template indiscriminately when encountering something. Rather, the agency-detector seems able to discriminate appropriately in many cases, but it likely features a *moderate* bias in the direction of allowing such false positives (Barrett, 2005, cf. further discussion in section 5). This bias towards assuming the agency of *conspecifics* relates to the asymmetry in cost between failing to detect conspecifics and failing to detect members of other species (see Duntley, 2005). In light of this, HADD is likely to feature a moderate anthropomorphizing bias.

Guthrie describes "literal anthropomorphization" as what occurs when we identify something non-human as a human. For example, one thinks one sees a human in the distance, but it turns out to be a tree. We usually correct such mistakes rapidly when we obtain more sensory input. We "partially anthropomorphize" when we see an event or a structure as having humanlike features or as involving human agency but without literally being a human. Finally, "fleeting anthropomorphism" occurs when we think we see a human face in something non-human like on the moon or in the clouds (Guthrie, 1995, 82-96).

According to Guthrie, religion is an example of these processes. It is anthropomorphism in the sense that it involves an attribution of human features to the non-human world. To believe in ghosts or in a deity is for example to anthropomorphize. This attribution is an erroneous inference, i.e. a false positive (Guthrie, 1995, 177-188). Guthrie argues that the sort of

---

notion that this involves a *hypersensitive* agency-detection device might need to be revised, they argue. Moreover, the finding that the bias does not increase in threatening situations seems to be at odds with the theory. One possible explanation of this finding, which Maij et al. (2019, 46) point to, is that the manipulations were not sufficiently realistic to trigger HADD. Maij. et al describe their conclusions as preliminary and point to the need for further investigation (2019, 49). In light of the preliminary nature of the conclusions, and the possibility that the manipulations failed to be sufficiently realistic to trigger HADD, we can let it remain an open question whether they provide grounds to revise HADD-theory.

anthropomorphism we refer to as religion is more systematic and resilient than similar inferences in other contexts, which tend to be literal, partial, or fleeting. For example, when mistaking a tree for a human, we rapidly correct the mistake as our sensory input increases in detail. By contrast, when subjects form a belief in a deity, this belief might persist for the rest of their life (Guthrie, 1995, 112, 200-201). In Guthrie's theory, this belief thus seems to be special in virtue of being a sort of perpetual anthropomorphism. Guthrie also provides many examples of anthropomorphism *occurring in religion*.[34] One example of this is that religious believers tend to attribute human features to deities to a much greater extent than what official religious doctrines usually warrant. Another example is that philosophers and theologians for a long time have struggled with making sense of deities, such as the God of Abrahamic theism, in ways that manage both to take account of his assumed transcendence and otherness and to present him in ways that are comprehensible to the human mind. The latter is hard to achieve without the usage of anthropomorphic themes (Guthrie, 1995, 179-185).

Importantly, as Guthrie points out, the human template involved in anthropomorphic thought specifies that humans have the capacity for language-based communication. Believers therefore tacitly extend this assumed capacity also to deities. They represent gods as humanlike agents with whom they may communicate and interact in meaningful ways. Subjects involved in religious rituals for example assume that the relevant deity will understand the rituals and respond to them. Subjects who pray assume that the deity to which they pray is capable of understanding them. Thus, they extend assumptions about human language-capacity to gods, spirits, and ghosts (Guthrie, 1995, 197-199). This theme, which we will encounter again in the work of Pascal Boyer and Lee Kirkpatrick below, is central to Guthrie:

> "The anthropomorphism of assuming language in gods appears ineluctable: there is no religion without relationship, no relationship without significant communication, no significant communication without language, and no language without likeness." (Guthrie, 1995, 199).

We assume that gods are agents with whom we can communicate in ways that resemble how we communicate with each other. Religion is thus a sort of *social relationship*, Guthrie importantly contends (1995, 202). Guthrie points out an important consequence of the notion that religion involves a social (personal) relationship. For someone to doubt a theistic doctrine

[34] This is not the same as Guthrie's primary claim, *that religion is anthropomorphism*. As Guthrie memorably puts it, "For many people, religious anthropomorphism consists of seeing God or gods as humanlike. In contrast, my claim is that God or gods consist in seeing the world as humanlike" (Guthrie, 1995, 178).

is not only to doubt the factual claim *that a god exists.* It is not akin to what a scientist does when realizing that the balance of the evidence favours scepticism with respect to his or her favoured theory. To doubt a theistic doctrine or teaching is to *undermine a personal relationship to a god* (Guthrie, 1995, 35-36, 202-203).

These important points indicate that religion involves much more than agency-detection. We have seen that agency-detection, the process HADD carries out, is the process by which we infer the presence of a living intentional being and make some preliminary discriminations about its nature (for example whether it is a conspecific). If HADD categorizes it as conspecific, then a barrage of further thought is to be expected, though not produced by the agency-detector. Instead, TOM is then triggered (Barrett, 2004). This indicates that HADD's inferences are too meagre to produce specific religious ideas about humanlike persons or to give rise to a social relationship to gods. Guthrie agrees, pointing out that the HADD-thesis seems "unnecessarily modest" and emphasizing the relative richness of the notion of a person compared with that of a mere agent (Guthrie, 2017, 77). Religious beliefs about spirits, ghosts, and deities do not arise in our minds as occasional false positives or by-products of the workings of our agency-detector. According to Guthrie (2017), we need the more complex notion of *anthropomorphism* to account for this.

It is not clear, however, that the notion of anthropomorphization is complex enough to capture the content of the often quite specific religious beliefs people have. Though anthropomorphism involves a template which is richer in content than the one animism involves, the former is nevertheless also a false positive, which occurs as the result of the mere natural functioning of the mind. The result of this is probably not beliefs about specific deities. But it is primarily the latter beliefs that religious subjects have. Religious subjects tend to believe in specific, personalized deities particular to their context rather than in ambiguous, non-identified, person-like entities. Also, the complexity of the many thoughts that subjects have about such agents when having already formed belief in them seems to require a richer and more complex psychology. Justin Barrett, Pascal Boyer, and Lee Kirkpatrick are therefore correct in introducing further evolved systems to account for the complexity of religious thought. Barrett (2004, 32) for example points out that the intuitions HADD produces tend to lead to further inferences about the desires and beliefs of the agent, and these are produced by TOM. Kirkpatrick argues (see also Boyer, 1994, 2002, and Barrett, 2004) that our minds comprise distinct evolved systems that govern agency-detection as well as the navigation of various aspects of *human social life*. The latter includes tasks such as forming and maintaining relationships between caregivers and offspring, handling social exchange, or navigating inter-

group interaction. According to Kirkpatrick, the complexity of our relationships to and interactions with deities reflects the many functionally distinct ways in which we engage each other (Kirkpatrick, 2005, 240-241). For example, subjects do not only (or even primarily) believe *that* deities exist. They engage deities as social agents who care about what we do and who might expect things from us. Furthermore, and to anticipate something we shall see later in this chapter, theists in the Abrahamic traditions do not typically engage God in brief encounters (akin to social exchange encounters for example). Rather, they in many cases form *lasting relationships* to God. In light of our analysis thus far then, this feature of religion is neither sufficiently accounted for by the spontaneous anthropomorphic inferences Guthrie describes nor by the intuitions HADD according to Barrett generates in ambiguous situations.[35]

Instead of explaining the complexity and long-lastingness of many religious beliefs by seeing it as reflecting a sort of perpetual anthropomorphism, we have, as Barrett (2004), Kirkpatrick (2005), and Boyer (2002) argue, grounds to see it as the result of how a larger collection of distinct processes solve various tasks. Kirkpatrick (2005) in particular emphasizes *attachment* and argues that (Abrahamic) religious belief in typical cases involves a strong emotional bond akin to the one that occurs in the relationship children have to parents. For a theist to doubt theism is, as Guthrie correctly points out, to undermine a social relationship. Given Kirkpatrick's theory, we can better account for why it is hard; to doubt theism is hard because it leads to fear of separation from an *attachment-figure.* But given Guthrie's (1995) thesis about our tendency to anthropomorphize, we can better understand why religious notions such as that of a god tend to be so anthropomorphic even though religious teachings often describe the gods in much more abstract ways. This tendency to anthropomorphize gods, which has been experimentally demonstrated by Justin Barrett and Frank Keil (1996), reflects our underlying, moderately strong proclivity to attribute human features to the natural world.

However, as an account of the formation of specific religious beliefs in subjects, even this complex picture does not seem sufficient. We also, as Helen De Cruz and Johan De Smedt (2015, 38) point out, need to introduce the notion of *cultural exposure*,[36] since religious beliefs do not spontaneously form from the mere functioning of our minds.[37] Pascal Boyer (2002)

[35] I therefore agree with Neil Van Leeuwen and Michiel van Elk, who contend that the relationship between the perceptual processes Guthrie (1995) describes and the sort of long-lasting religious beliefs subjects tend to have is not explained sufficiently by the anthropomorphization-thesis (Van Leeuwen and Van Elk, 2018, 3).

[36] We shall look at this topic in more detail in section 6, where we look at a proposal Gervais et al. (2011) defend. We again return to it in chapter 3, section 3, where we look at work by Lanman and Barrett (2008). It is also relevant in our present discussion of the role of anthropomorphism and HADD in the formation of religious beliefs.

[37] That full-blown religious beliefs arise directly from the mere functioning of cognitive systems, like HADD, is not something advocates of HADD theory claim. For example, Barrett (2004) clearly recognizes that there is an interplay between cultural input and cognitive biases, and this is highlighted even more by Barrett and Lanman

makes a similar point: to account for the formation of religious beliefs, we need to assume the already established presence of religious ideas in the surrounding culture. Boyer argues that when people talk and thus expose each other to such ideas, we plausibly can expect an activation of many of the above-specified systems. Then, their activation can give rise to religious thought (Boyer, 2002, 168).

This need for cultural exposure, which introduces a distinct religious concept to the mind, applies when we ask the *proximate* question of how religious beliefs arise. Proximate questions are, roughly, about the "nuts and bolts" of the mechanisms and processes involved in cognition. *Ultimate* questions, by contrast, concern issues like the evolutionary origin and adaptive function of a given trait, mechanism, or process (see Ch. 3, section 2 for further discussion). CSR is strongly interested in the ultimate origin of religious concepts and aims to account for it by studying how cognitive biases make us susceptible to religious notions and constrain their subsequent cultural evolution. The MCI theory, which we look at in section 5, is a central attempt to account for this. We shall see that the ultimate question of how religious concepts and beliefs arise and propagate can be answered in terms of human cognition. But when it comes to the proximate origin of religious beliefs in subjects presently, we must, as mentioned, also introduce cultural exposure to such concepts as an explanatory factor. A central topic in the next chapter is how work in CSR on how general processes and cultural input interact sheds important light on proximate questions, including belief-formation. For now, to get a taste of this, let us briefly appreciate how HADD is relevant to both ultimate and proximate issues.

HADD seems to play a role in answering both ultimate and proximate questions concerning religion. Regarding the former issue, Barrett (2004, 33) points to the possibility that HADD can contribute by providing agency-intuitions in ambiguous situations where no agent actually is present. Such intuitions might then form the initial basis for religious ideas which subsequently propagate in culture, thereby developing into more elaborate forms. Justin Barrett and Jonathan Lanman (2008) further point to how the joint operation of HADD and TOM makes notions of intentional agents highly relevant to us. This relevance in extension also applies to supernatural agents, making them likely candidates for such cultural propagation. HADD and TOM thus help sustain this propagation-process (Barrett and Lanman, 2008, 114-116, cf. the discussion of MCI-theory in section 5, which points to further factors that sustain and influence this process).

(2008). But as Van Leeuwen and Van Elk (2018, 5) point out, some simplified, popularized depictions of HADD-theory seem to suggest this. It is therefore worth pointing it out.

Can HADD contribute proximately to the production of religious beliefs? This also seems to be the case (as we shall see in more detail in the next chapter). As J. Barrett (2004) argues, HADD-generated intuitions can together with cultural input produce beliefs about the involvement of supernatural agency in a concrete situation (see Ch. 3, section 5 for details). In a continuation of this view, Barrett and Lanman (2008) focus on the interplay between cultural input and situationally specific agency-intuitions, suggesting that HADD can contribute to explain how some specific beliefs arise in subjects who have been exposed to religious concepts and who already have religious belief (Barrett and Lanman, 2008, 116). Finally, consider Neil Van Leeuwen and Michiel Van Elk (2018), who offer a critical appraisal of HADD-theory. Van Leeuwen and Van Elk are critical of the notion that HADD directly causes what they refer to as "general religious beliefs", meaning beliefs in the existence of a given supernatural agent. Instead, cultural learning is a much more important explanatory factor regarding such beliefs (Van Leeuwen and Van Elk, 2018, 7-9).[38] Nevertheless, Van Leeuwen and Van Elk also argue that agency-related intuitions together with general religious beliefs can lead to the formation of "personal religious beliefs", meaning religious beliefs that are about the person who has the belief (2018, 13-18). These examples illustrate that HADD is relevant to religious belief-formation even though this system alone (or even together with TOM, the attachment-system, or systems for social exchange) does not fully explain the formation of religious beliefs. Van Leeuwen, Van Elk, Boyer, De Cruz, De Smedt, Barrett, and Lanman seem right: we also need to introduce cultural input to proximately explain how beliefs form. We continue our exploration of this topic in section 6 and in more detail in chapter 3.

In conclusion, based on our discussion in the present section, Guthrie's theory of anthropomorphism, Justin Barrett's HADD-theory and its current applications and developments plausibly contribute to explaining how religious ideas ultimately have arisen as by-products of how the mind solves fitness-relevant tasks, and the interplay of HADD and such ideas helps explain, proximately, how religious beliefs arise in concrete situations presently.

Another factor (besides HADD) Barrett (2004) and Barrett and Lanman (2008) point to as important in explaining the propagation and formation of religious beliefs is *the cognitive attractiveness* of religious concepts. The next two sections describe such concepts and discuss what makes them cognitively attractive. I believe the account Pascal Boyer (2002) provides of

[38] Van Leeuwen and Van Elk suggest that the evidence therefore contradicts "strong versions" of HADD theory according to which agency-detection directly causes religious beliefs (Van Leeuwen and Van Elk, 2018, 10). But given our analysis of Barrett's (2004) HADD-theory above and given Barrett and Lanman's (2008) explicit recognition that cultural input matters to the formation of religious beliefs, it is not entirely clear whether this criticism applies to any representative, current version of HADD theory.

these concepts is still the most compelling and we therefore base the following analysis on it, before introducing the related "MCI theory" and some recent discussions of it in section 5.

4. Pascal Boyer and the structure of religious concepts

Pascal Boyer is (like Guthrie) an anthropologist who started developing a cognitive approach to religion early (1994), and whose work is influential in CSR. The present section is about Boyer's (2002) account of the structure of religious concepts and the distinction he makes between concepts and underlying templates consisting of naïve, ontological intuitions. According to this account, which Barrett (2004) and for example McCauley (2013) develop further, religious concepts are largely intuitive because our own minds provide intuitions and assumptions that become part of their content. However, these concepts are also moderately counterintuitive because some of their content, especially content obtained through social interactions, violates these intuitions and assumptions. They are therefore, as Justin Barrett (2000) puts it, "minimally counterintuitive" (MCI) (see section 5 for a discussion of MCI theory).

A central idea in Boyer's theory (1994, 2002) of religion is that this sort of combination of intuitions with specific information is *cognitively attractive* and therefore likely to propagate. The work of the anthropologist Dan Sperber on cultural evolution lays an important foundation for this theory. Sperber (1996) argues that to explain culture is to explain why some concepts (i.e. mental representations) become widespread and persist in human populations over time. Such representations are not only private things in our heads but cultural in the sense that they occur in similar forms with a high frequency in human populations. Such an explanation must give an account of why some concepts rather than others are successful. Sperber calls the explanation he proposes an "epidemiology". This means an explanation of the success some types of concepts have in terms of cognitive propensities that favor them. Some concepts are "contagious", which means that our minds easily construct them from information our surroundings provide. Moreover, they are memorable, readily processable in our minds, and easy to transmit. According to Sperber, we are susceptible to such concepts analogously to how our bodies are susceptible to viruses (1996, 49-50, 57-61). Pascal Boyer (2002) builds on this account of culture in his own study of religion. We can think of his theory of religion as an

epidemiology of belief applied to that subset of cultural representations which involves purported supernatural agents. According to Boyer, to explain religion is to explain a mental epidemic involving the propagation of these representations. Such an explanation must account for why these rather than other concepts are contagious (Boyer, 2002, 53-54). Let us now investigate the structure of the religious concepts Sperber, Boyer, and Barrett see as contagious.

We start by appreciating that Pascal Boyer agrees with Guthrie that religious concepts are about humanlike agents and that religion involves a relationship with such supposed agents. Moreover, religious concepts (of the kind likely to propagate culturally) are to a large extent intuitive. However, Boyer (2002, cf. McCauley, 2013, 162-171) also suggests that these concepts contain information that *violates* tacit expectations we have about reality, making them to some extent *counterintuitive.* The agents of religious belief are humanlike but also different from us in ways that make them interesting and memorable. This is something Guthrie (1995, 2017) strongly disagrees with.[39] Boyer thus attributes two opposing features to religious concepts:

(1) Religious concepts connect with "ontological domains" and are largely *intuitive.*
(2) Religious concepts violate expectations pertaining to the domain to which our minds relate them and therefore are *counterintuitive.*

Let us start with (1). An "ontological domain" is an assumed "compartment" of reality that our minds automatically carve as the result of how various psychological systems solve problems. Distinct intuitive principles and heuristics pertain to distinct ontological domains (Boyer and Barrett, 2005, 96). Boyer suggests the domains *person, animal, plant, artefact,* and *natural object*. Sets of intuitive expectations that constitute naïve theories or ontologies pertain to these domains. These are *naïve physics, naïve biology, and naïve psychology* (Boyer, 2000, 196-198, 2002, cf. Barrett, 2008a, which makes some adjustments to this picture). These domains do not necessarily match real ontological categories but only how our minds carve reality. Furthermore, as we shall see, the naïve theories related to each domain sometimes overlap. This is for example the case with person and animal, which both involve intuitive physics, biology, and intuitive attributions of some kind of intentionality.

The domain for *natural object* "activates" naïve physics. This means that when a subject considers an object identified as belonging to this domain, then the subject's mind automatically

[39] The final part of section 5 in the present chapter deals with Guthrie's criticism.

generates intuitive beliefs constitutive of naïve physics. Naïve physics arises in infancy as part of normal cognitive development. It consists of beliefs about how physical objects behave. For example, we expect that such objects do not pop in and out of existence and that two objects that collide do not merge into one object (Boyer, 2000, see also Barrett, 2004, 4, Spelke and Kinzler, 2007, McCauley, 2013, 61-70).

The domain for *animals* activates *naïve biology* and naïve physics. This means larger sets of beliefs automatically apply. Naïve biology consists of tacit beliefs according to which entities can initiate goal-directed movement. Naïve biology also consists of biological essentialism. This for example includes the assumption that an animal of one species cannot produce offspring belonging to another. This means that cosmetically changing the appearance of an animal does not change what kind of animal it is. Biological essentialism involves the assumption that each species has an inner essence or core that determines an unchangeable nature (Gelman, Coley, and Gottfried, 1994, Boyer, 2000, Atran, 2002, Barrett, 2004, Boyer and Barrett, 2005).

The domain for *person* activates *naïve psychology*, naïve biology, and naïve physics. Naïve psychology consists of expectations we intuitively have towards entities we represent as *persons*. These expectations constitute a *Theory of Mind* (TOM). We employ TOM when interpreting the behaviour of agents by attributing beliefs and desires. This is what happens when HADD (i.e. the "categorizer") has judged that the agent is a conspecific (Barrett, 2004). TOM is according to Boyer and Barrett (2005) to some extent also involved in interpretation of animal behaviour, which involves the attribution of goal-directedness and intentionality. However, as Boyer and Barrett (2005) further argue, the intuitions and expectations that occur when we respond to human behaviour involve a richer psychology that relate to a greater number of distinct types of interactions humans are capable of, such as social exchange, inter-group conflict, interactions with kin or mating and courtship. This suggests that at least some distinct aspects of TOM pertain only to agents that we identify as human persons (i.e. conspecifics) (Boyer, 2000, Boyer and Barrett, 2005, 105-106, Barrett, 2005, McCauley, 2013, 76-82).

The domain for artefacts triggers teleological thinking. Encounters with objects that our minds identify as artefacts trigger spontaneous questions about design and purpose. Such encounters compel us to think about what the thing is for or who made it (Boyer, 2002, 116, Boyer and Barrett, 2005, 102-104). Teleological thinking allows subjects easily to learn how to use tools by inferring their intended purposes. Being prone to apply such thinking arguably conferred great advantages to our ancestors. Encounters with non-artefacts that seem designed

can also trigger teleological thinking. When this occurs, the result is a false positive, that is, we falsely infer a purpose or causal relationship when we in fact have encountered a largely random phenomenon (Dawkins, 1998). There is evidence for a bias towards teleological reasoning in human thought. Young children for example exhibit a tendency to think of non-artefacts (natural structures like rocks for example) as having been designed and as having purposes. Deborah Kelemen labels this effect "promiscuous teleology" (Kelemen, 2004, 295-299).

With this account of ontological domains in mind, let us turn to what it means for a concept to be largely *intuitive.* What makes a concept intuitive is that the entity it purports to describe and the features it attributes to this entity match the naïve theory or theories that the concept activates (Boyer, 2000, 197, 2002, 58-105). For example, the concept of a human person is intuitive because it describes an entity we represent as a person and the concept does not contain information that violates the expectations relating to naïve psychology, biology or physics. Boyer (2002) contends that religious concepts are *largely intuitive*. For example, theistic concepts purport to describe humanlike agents. This means naïve psychology applies. The difference is that the concept specifies additional properties that violate naïve biology and physics. That is why the concept is *largely* rather than completely intuitive. I return to this below. Boyer argues that what primarily establishes the intuitiveness of the concept is *information our own minds provide*. For example, information from social surroundings is not the source of the beliefs that a god knows things or has opinions about human affairs. Our minds assume this in virtue of identifying this god as a person (Boyer, 2002).

In order to illustrate how the mind contributes to the content of concepts, Boyer presents an account of concept-acquisition involving a distinction between concepts and templates. A basic assumption that guides this account is that we build concepts from incomplete information that our social and cultural context provide us. Our minds fill out details and make spontaneous further inferences thereby making the concept complete (Boyer, 2002, 46-47). Boyer asks us to consider a teacher who shows a child a picture of a walrus and provides information about these animals. When this happens, the child intuitively categorizes the information about walruses as an entry under an ontological domain for animals. Naïve physics and naïve biology now apply, and the mind makes use of a tacit systematic mental arrangement (template) to make sense of the intuitions and assumptions this gives rise to. The animal template includes information that the child expects to apply to *all* animals. It therefore enables the child to make sense of the information the teacher provides about walruses, which eventually becomes a distinct representation subsumed as an instance of animals. We here have two concepts with different levels of generality, "walrus" and "animal". When the child learns the former concept,

the information specified by the template for animals automatically applies to walruses. This template-internal information constitutes a significant part of the content of the concept of walrus. For example, the child tacitly applies the biological essentialism that the animal-concept involves. That walruses give birth only to walruses is not something the child must learn. The child automatically believes this. The concept also includes information about walruses that the child must learn. This includes information about how they look or what they eat. This information is not an inherent part of the concept of animal. It is information that the social context provides (Boyer, 2000, 197-198, 2002, 48-51).

As long as socially transmitted information does not suggest features that violate the naïve ontological theories mentioned above, also such information accounts for the intuitiveness of the concept. We can say that this intuitiveness is dependent on template-external (i.e. socially transmitted) information. The alternative is intuitiveness which is dependent on template-internal information, i.e. assumptions the mind itself automatically provides.

The learning of a religious concept resembles what occurs in the above example but with some important differences. Imagine a child who learns a theistic concept from his dad in childhood. He listens to his dad's stories about a distinct supposed deity, a powerful person who created the whole world and who is aware of everything we do and think. Dad tells the child that God punishes evildoers and that humans have inherited something called "sin". Luckily, God had a son named Jesus who knew this and chose to take the punishment we deserve upon himself by dying on a cross. He subsequently arose from the dead and went to heaven. We can pray to him and he will listen. If we do something bad, then we should tell him we are sorry. What occurs during this learning-episode is to some extent similar to what occurred when the child learned about walruses. But let us start with how this episode is different from the one that involved walruses.

In the walrus-example Boyer provides, the subject obtains the concept "walrus". This is a representation that allows him to categorize information about individual walruses. This differs from what happens when the child obtains the theistic concept "God". Learning a theistic concept means learning about a specific purported person one calls "God" (or Zeus, Baal, and so on). A theistic concept is thus not analogous to the concept "walrus". The latter but not the former is a category-concept. However, "God" can also be used as a category-concept.[40] This concept of "God" allows the child to categorize different individual purported deities, one of which he might nevertheless address personally in worship, using "God" as a proper name. In

[40] It is then not used the way we employ the term in the present investigation, for the deity of Abrahamic theism.

the episode above, the child does not acquire the concept "God" in the category-sense but a specific theistic concept for a specific purported person. Eventually he might nevertheless learn to fluctuate between using "God" in these two ways. Let us now turn to a similarity.

It follows from the above account of naïve ontology that the walrus-episode and the theistic learning episode are similar in an important respect. The theistic concept and the walrus-concept have a dual structure. They both contain intuitions that our own minds provide, and they contain further information provided by the social context. Thus, in both cases the mind informs the concepts to a large extent. Consider again the story about God listening to our prayers and caring about us. This indicates that God is a person. The representation therefore contains assumptions based on naïve psychology. The information in this case is systematized under the template that arranges information about persons. Being able to listen to prayer indicates psychological abilities that we do not assume animals have. As Boyer contends, when we represent something as a person, expectations we automatically have towards persons apply (2000, 197). This includes (as Guthrie argued) the expectation that this is an entity with whom we may communicate using human language. Here we have a match between intuitive expectations ("persons have psychological states") and information provided by the surroundings ("God loves us and listens to our prayers"). The concept thus combines what we can call template internal and template external intuitiveness.

Let us now turn to the other feature Boyer (2002) attributes to religious concepts. They describe features that *violate* some of the naïve expectations that pertain to the ontological domain to which our minds relate them. Consider for example the notion that God is an immaterial spirit. Because this alleged feature violates naïve physics, it is not intuitive. Furthermore, though large parts of the story that dad told the child (in the example above) are intuitive, the notion that Jesus rose (physically) from the dead is *not* intuitive. It violates naïve biology since it suggests that bodily functions continue after death. Such information makes the concept of God or Jesus *counterintuitive*, but not strongly so (given that many default-assumptions continue to apply). Here the theistic concept differs from the concepts of a human being or a walrus because the two latter do not violate intuitive ontologies. The violation in the case of an intentional being who lacks body is what Pascal Boyer calls a *breach*. It is a breach with naïve biology and physics. Moreover, the notion in theism of an omniscient God (a deity who knows everything it is logically possible to know) involves a breach with intuitions involved in naïve psychology according to which agents have limited knowledge.[41] Another

[41] As we will see in chapter 5, Rebekah Richert and Justin Barrett challenge the claim that thinking of deities as having unlimited knowledge is counterintuitive. I will there defend this claim from their challenge.

type of violation is a *transfer*. A transfer occurs when we relate a feature that belongs in one domain to an object that intuitively belongs in another domain. For example, the features of having beliefs and desires belong to objects we identify as persons. If we relate those features to an artefact, then we transfer them to an object that belongs to another domain than person. This makes the resulting concept counterintuitive. An example of this is the notion of a statue that listens to prayers, or a table which is happy when people have dinner at it (Boyer, 2000, 2002, 71-84, cf. McCauley, 2013). That a concept is counterintuitive in this sense does not, as Barrett and Nyhof point out, mean that it is difficult to process or hard to accept. It means that the concept contains information which is at odds with naïve ontological intuitions pertaining to the relevant category. In their view, this minor tension with our intuitions makes such a concept attractive (Barrett and Nyhof, 2001, 72).

Boyer (2000) presents a taxonomy of 15 kinds of concepts that involve various combinations of breaches and transfers for the five ontological domains specified above. He argues that this taxonomy captures important features of the religious concepts that are widespread in human cultures (see also Boyer and Ramble, 2001, 536). This means that what at first seems to be a set of highly variegated cultural expressions can be broken down to a limited number of structures. The central theme that recurs in these structures is that of a humanlike agent with whom we can communicate. We now need to ask why concepts with this specific structure are, as Dan Sperber (1996) puts it, "contagious". We now turn to MCI-theory, a central attempt in CSR to answer this question.

5. The memorability and relevance of religious concepts: the MCI theory

A central thesis in CSR is that all over the world, subjects widely entertain concepts of slightly counterintuitive intentional agents with whom one may communicate in human ways. This is alleged to be a central, cross-culturally stable feature of religious traditions both past and present, which occurs in both small-scale groups and in large-scale societies (Boyer, 2002, 58-105, Atran, 2002, 93-100, Atran and Norenzayan, 2004, Norenzayan, 2013). To explain this pattern, Justin Barrett, Pascal Boyer, and Lee Kirkpatrick (among others) argue that we can expect intentional agent concepts with a limited number of violations of naïve ontological assumptions to propagate because such a structure is *cognitively attractive* compared with other structures (Boyer, 1994, 2000, 2002, 2018, Barrett, 2000, 2004, Barrett, 2008b, Barrett and

Lanman, 2008, Kirkpatrick, 2005, McCauley, 2013). As mentioned, J. Barrett (2000, 30) calls this structure *minimal counterintuitiveness* (MCI) and the concepts that have it "MCI-concepts". According to Barrett, the central aspect of MCI concepts which makes them attractive is that, to a limited extent, they are *counterintuitive and therefore memorable.* An additional property which makes a subset of such concepts especially attractive is that they describe *intentional agents* with access to information that matters to us in social interactions (Barrett, 2008b, 150-154). As a result of how we evolved in social groups and therefore became highly sensitive to information relevant to social interaction, this subset has high "inferential potential" (Boyer, 2002). This makes such concepts *psychologically relevant* (in a specific sense we shall discuss). *Memorability*, brought about by minimal counterintuitiveness, and *relevance*, brought about by strong inferential potential, are the two central properties which make religious concepts cognitively attractive and therefore likely to propagate. We can refer to theories that appeal to the cognitive attractiveness of this structure in an explanation of the successful cultural propagation of religious concepts as MCI theory.

We start our discussion of MCI theory with a look at the thesis that MCIs are memorable. This thesis has been extensively studied, and it has received significant empirical support. Inspired by F. Bartlett's (1932) early work on memory, Justin Barrett and Melanie Nyhof (2001) carried out 4 experiments to test the claim that MCI concepts, which they describe as cross-culturally prevalent, have transmission-advantages in virtue of being particularly memorable. They were especially interested in how this cross-cultural prevalence or trend can be explained in terms of how universally shared mechanisms interact with cultural material in particular contexts (Barrett and Nyhof, 2001, 70). Barrett and Nyhof point out that much previous research on recall has focused on recall of information embedded in *schemes*, that is, culturally specific representational structures, obtained through experiences, by which subjects make mutual assumptions about each other's knowledge in ways that enhance their communication (Barrett and Nyhof, 2001, 70-71). According to Barrett and Nyhof, since schemes are culturally specific and based on particular experiences, they are not good candidates for an explanation of *cross-cultural* trends in how ideas propagate. Instead, Barrett and Nyhof argue, we should focus on Boyer's (1994) theory of how religious concepts violate representational structures that do not depend on culturally specific input but naïve ontologies which arise automatically in terms of the natural functioning of the human mind (Barrett and Nyhof, 2001, 71-72). In one of their experiments, the subjects were exposed to native American folktales and then, after having done a distractor-task, they had to retell them. To control for cultural familiarity as a potential confounding factor which could affect the recall-rate, Barrett

and Nyhof made sure none of their participants were native American. They found better recall and more precise retelling of the elements containing counterintuitive concepts than intuitive concepts (Barrett and Nyhof, 2001, 75-77). In this and the 3 other experiments, Barrett and Nyhof found that concepts with one violation of naive ontological intuitions have transmission-advantages compared to purely intuitive concepts, thus supporting Boyer's theory (Barrett and Nyhof, 2001, 91). They further suggest (pp. 93-94) that Boyer has identified a "cognitive optimum" for religious concepts, a structure with the right balance of intuitive and attention-catching properties to make us susceptible to it.

A range of studies have found similar effects. For example, Boyer and Ramble (2001) found recall-advantages for concepts containing violations of naïve ontological assumptions, a finding they made with subjects in multiple culturally distinct settings (Gabon, France, and Nepal). In a more recent study, Banerjee, Haque, and Spelke (2013) were interested in whether the bias towards MCI concepts these earlier studies revealed is present in children too. They found evidence for this. Children between 7 and 9 years of age showed better recall for concepts containing a small number of violations of naïve ontology than concepts not containing such violations. Moreover, consistent with Boyer's cognitive optimum idea, recall for concepts with too many ontological violations was not better than recall for intuitive concepts. Harmon-Vukic, Upal, and Sheehan (2012) also found that MCI concepts have memory-related transmission advantages compared to intuitive and strongly counterintuitive concepts, thus also supporting the notion of a cognitive optimum. By contrast, Porubanova-Norquist, Shaw, and Xygalatas (2013) found a different pattern. They found evidence for stronger immediate recall for concepts that violate *culturally learned expectations* than concepts that violate *naïve ontological intuitions*. However, they also found stronger recall for the latter concepts than intuitive concepts after a delay. In a discussion of these findings, Porubanova-Norquist, Shaw, and Xygalatas suggest that culturally unexpected information has the best chances of being recalled, whilst MCI concepts nevertheless exhibit long term memory-advantages compared to intuitive concepts (Porubanova-Norquist, Shaw, and Xygalatas, 2013, 188). These findings point to a more complicated pattern than initial versions of MCI-theory, but they still support the central thesis that MCIs have transmission-advantages. As a final example, Scott Atran and colleagues found that subjects had better (immediate) recall for intuitive concepts than MCI concepts and concepts that violate many ontological assumptions. Nevertheless, MCI concepts degraded less rapidly than intuitive concepts, which in Atran's view suggests that they are "cognitively resilient" (Atran, 2002, 105). Moreover, Atran and colleagues also found that recall of sets of beliefs containing a small proportion of MCI concepts was better than recall of

completely intuitive or massively counterintuitive sets. This indicates that injecting a modest proportion of MCIs into largely intuitive narratives, something Atran sees evidence of in the stories and tales of religion and mythology, makes these narratives cognitively attractive (Atran (2002, 106-107). The evidence thus indicates that MCI-concepts are memorable compared to other concepts, especially when we consider long-term memory. In conclusion then, the first component of MCI theory, the memorability-thesis, has empirical support.

The other central part of the cognitive attractiveness of MCI concepts is their *psychological relevance*. This notion relates to the work of Dan Sperber and Deidre Wilson on communication. According to Sperber and Wilson (2004), relevance is a property of input to cognitive processes. Such input can be external in the sense that it is a perceptual experience of an object in our surroundings or the hearing of an utterance. Such input is internal if it is a thought, a mental image, or a concept that the mind activates. The relevance of input to a process is context-relative. Input the processing of which contributes to the handling of a specific challenge in a specific context is relevant in that context. Furthermore, the less cognitive effort necessary to process the input, the more relevant the input is. Input is relevant if it connects with representations already established in the mind and leads subjects easily to infer many new things in ways that help them solve problems (Sperber and Wilson, 2004, 608-609).[42]

Boyer (2002) builds on this account but also uses the term "relevance" for a further property: the ability of a concept to trigger the functioning of many rather than few evolved cognitive systems. This notion of relevance indicates a stable feature concepts can have in relation to species-typical, evolved features of human minds. Boyer calls it "aggregate relevance" (AR) (2002, 56). Concepts with high AR have the ability to trigger or activate many different evolved cognitive systems rather than few. For example, concepts of intentional agents have high AR because of the many distinct systems involved in handling human social interaction. Concepts with high AR tend to lead subjects to make many inferences and form many beliefs easily compared with concepts with low AR (Boyer, 2002, 56-57). Such concepts are able to latch on to more proximate cognitive events and intuitive processes related to social interaction in a wide range of situations, thus readily establishing themselves and enduring in our minds. Their high relevance contributes to making them sticky and hard to get rid of. Concepts of spirits, demons, and gods have high AR. These concepts connect with output from HADD (Barrett, 2000, 2004), but are also good at latching on to processing occurring in social exchange situations (Boyer, 2002), as well as systems devoted to attachment (Kirkpatrick,

[42] For a detailed presentation of relevance-theory, see Sperber and Wilson (2004).

2005). These concepts then provide endless opportunities for further thought. This is because subjects do much more than merely inferring the existence of counterintuitive agents in general or their presence in concrete situations. They also form thoughts about what such agents think about what is going on in the situation and what they intend to do next (Boyer, 2002, see also Barrett, 2004, Barrett and Lanman, 2008).

According to Barrett, a mere counterintuitive feature makes a concept memorable, but memorable concepts of *intentional agents* are also psychologically relevant. Given how sensitive we are to information related to social interactions, such concepts have "inferential potential". Therefore, if we add this property to an MCI concept, it becomes a much better candidate for a god-concept (Barrett, 2008b, 152, Barrett and Lanman, 2008, 114-115, Boyer, 2002). As mentioned above, this is also what scholars find: religion consists of concepts of *agents* rather than just weird stuff we cannot interact meaningfully with. As a contrast to concepts with inferential potential, Boyer describes the concept of a "divine brute", a deity *without* any conscious states. Why do we not find such a concept in religion? It is for example counterintuitive in virtue of a breach with physics. Why does such a concept fail to get itself embedded in many minds? Boyer's answer is that it has very low AR. Very little of the practical, intuitive thought that occurs in our daily lives would be able to connect with this concept. It therefore has very low inferential potential (Boyer, 2002, 188, see also Barrett, 2008b).

In summary thus far, the relevance of concepts pertains to how effortlessly their activation allows us to produce many new thoughts and to what extent they trigger many rather than few evolved systems. Since we are so sensitive to socially relevant information, MCIs of agents are especially relevant. Together with the memorability-thesis, this account of relevance helps us see why the sort of concepts we tend to find in religion are cognitively attractive and therefore likely to propagate. The next section looks closer at why culturally distinct versions of such attractive concepts congregate and stabilize in distinct cultural contexts. Moreover, and this relates to our goal in the next chapter of exploring the formation of religious beliefs, we shall explore how humans use the beliefs of others in the group to which they belong as intuitive cues when forming their own beliefs. This tells us something important about why people are selective about what exact counterintuitive agents they believe in despite the fact that a wide range of them are memorable and relevant. However, before we get to that, let us observe that MCI-theory recently has been strongly criticized. We shall in the following turn the focus to Benjamim Grant Purzycki and Aiyana K. Willard (2015), who scrutinize MCI theory and offer some important criticisms. I shall also consider two possible objections based on criticism Guthrie (2017) offers. I argue that none of these provide grounds to reject MCI theory, even

though they involve relevant points. I end with some remarks about what implications, if any, the outcome of this debate has for a central argument I offer in chapter 7.

Purzycki and Willard (2015) reconstruct MCI theory as consisting of the following four claims. The first concerns what we discussed in the previous section, namely the existence of various domains of evolved, intuitive assumptions about distinct parts of reality. Secondly, we have the memorability-claim: MCI concepts are more memorable than intuitive or strongly counterintuitive concepts. Third, MCI advocates contend that religion consists to a large extent of MCI concepts. Finally, the cultural success of religious concepts can be accounted for, partly, in terms of their memorability (Purzucki and Willard, 2015, 1). Among the many issues Purzucki and Willard (2015) deal with and the many relevant points they make, I shall focus on two central worries. The first is about what Purzycki and Willard see as conceptual confusion in MCI theory concerning what an MCI is and how we can distinguish it from other concepts, in particular, those that are not counterintuitive but nevertheless are surprising, unexpected, or culturally anomalous. The second issue concerns the evidence for the specific ontological intuitions MCIs allegedly violate.

We start with the conceptual confusion problem. To see the problem Purzucki and Willard are getting at, we need to look at their distinction between what they call "deep inferences" and "shallow inferences". The former refers to the implicit and automatic intuitions evolved systems produce, and the latter refers to cognitively accessible, more reflective information about specific conceptual relationships. An example of the former is the intuitive inference that an object which self-propels in a seemingly nonrandom fashion has a mind, and an example of the latter is that a given object is a hungry dog that wants a bone (Purzycki and Willard, 2015, 6-7). The latter type of information is organized in schemes, defined as cognitively accessible structures that organize specific pieces of information based on experiences and ideas one is exposed to in one's context. Shared schemes are examples of cultural schemes (Purzycki and Willard, 2015, 9-10).

Now to the conceptual confusion problem. It is a crucial claim in MCI theory that MCI concepts violate *deep inferences*, not shallow ones. We saw for example how Barrett and Nyhof stress the importance of violations of universally shared intuitions rather than information organized in schemes. But according to Purzucki and Willard, MCI theory has a serious problem in virtue of being afflicted with a "chronic confusion" about the distinction between deep and shallow inferences (2015, 14). One reason this is problematic is that violation of schematic content also potentially affects recall (Purzycki and Willard, 2015, 15, cf. Barrett and Nyhof, 2001). This indicates a possible confounding factor at play. If schematic content affects

recall and MCI scholars do not control for this when exposing their subjects to MCIs, then it would be unclear whether the results were due to the effects of violation of schematic content or violation of universally shared ontological intuitions. Another reason conceptual confusion is problematic is that it might mislead scholars to construct quasi MCI-concepts which violate only shallow inferences. In that case, high recall would fail to provide evidence for the theory, since they would not in fact have tested recall of *MCI concepts* but another type of concept.

Are scholars positive to MCI theory conceptually confused concerning the difference between culturally specific schemes and general human ontological intuitions? Given how Barrett and Nyhof (2001) controlled for cultural familiarity when exposing their subjects to MCIs and given their discussion of how schemes are distinct from universal ontological intuitions, that study does not seem to involve confusion. But that does not mean such confusion does not appear in other studies or in MCI theory in general, as Purzycki and Willard's thorough review seems to indicate. Supposing there is a conceptual confusion problem in current MCI research, what does that tell us about the truth of MCI theory? It tells us that MCI scholars ought to be more conceptually precise and careful when carrying out studies. Such precision seems, as the example of Barrett and Nyhof (2001) illustrates, possible to achieve. Moreover, that Purzycki and Willard's criticism is intended in this way, as exhorting MCI-scholars to improve rather than to find a new vocation, is indicated by their helpful and relevant suggestions concerning how future MCI experiments can obtain improved designs (2015, 22-23). Presumably, one would not offer help to a theory found hopelessly implausible. We conclude, the conceptual confusion challenge, to the extent that it appears in current MCI-scholarship, can be met with improved experimental designs.

Let us now turn to the intuitive ontology objection. Similar worries appear in Guthrie's criticism, to be discussed below. Purzycki and Willard (2015) raise concerns about the content and origin of the intuitions MCI concepts purportedly violate. For example, in one study the concept, "virgin mother", is treated as an MCI concept, since it purportedly violates naïve ontological assumptions according to which pregnancy requires sexual activity. In another study, the concept of a "breathing table" is seen as MCI because it transfers "breathing", a property assumedly belonging to the Animal category, to that of Artefacts (Purzycki and Willard, 19-20). This is problematic, Purzycki and Willard argue, since there is no evidence that the naïve ontology for the Animal category consists of distinct intuitions about breathing or the need for sex to create babies (2015, 20). Moreover, it does not suffice to have evidence that subjects widely form such intuitions, since not all intuitions arise out of evolved cognitive systems. It is also possible that universal "folk-beliefs" about how things are in the world can

make some notions intuitive. Violations of folk-beliefs are not MCI (Purzycki and Willard, 2015, 7). Do MCI-scholars lack evidence for the ontological intuitions MCIs violate? A conceptual confusion problem also looms here. Such scholars might mistake these intuitions for schematic information or universal folk-beliefs. Let us explore this.

Recall Boyer's (2002) distinction between concepts and templates (from the previous section). It seems correct that the template for Animal does not contain specific intuitions about how animals breathe. This information seems more akin to conceptual, schematic information subjects obtain by talking to others, reading books, or watching television programs. We saw that Boyer (2002) distinguishes specific *concepts* like that of a walrus, which contains information about where walruses live or what they eat, and the underlying Animal *template*, which contains a species essentialism that automatically applies also to other animals. Intuitions or beliefs about how an animal breathes seem to be conceptual information, not template-internal intuitions. The problem of a lack of evidence for the relevant ontological intuition which Willard and Purzycki point to therefore seems real in this case. The problem could also reflect conceptual confusion. There is a danger that widespread, conceptually rich *template-external* information is mistaken for *template-internal* intuitions about all entities belonging to Animal. This can create errors in construction of MCI-items. For example, the concept of a walrus with properties that only violate template-external information is not MCI. An attempt to create an MCI by only violating information about the habitat, smell, breath, or idiosyncratic habits of walruses, would therefore fail.[43] Purzycki and Willard (2015) argue that MCI scholars frequently create such concepts when attempting to reveal an MCI effect (for examples which indicate this, see pp. 20-21). Given that this is a problem among MCI-scholars, Purzycki and Willard are therefore right to ask for clarity, care, and precision here. Moreover, those alleged MCIs based on violation of ontological intuitions whose existence lack evidence must also be removed when testing MCI-theory. In light of this, there appears to be some real issues and dangers in tests of MCI-theory, which scholars must address, again, with improved experimental designs, as well as a careful distinction between the domain of intuitive ontological intuitions (which MCIs violate) on the one hand, and schematic cultural information or universal folk-theories on the other. However, this does not show that MCI-theory must be rejected. Instead, it reveals some issues with how it is pursued and some pitfalls that must be avoided.

[43] By contrast, the concept of a walrus which gave birth to a dog is MCI, given how it violates the species essentialism the naïve ontology for animals involves.

Let us finally turn to two objections based on criticism Stewart Guthrie (2017) offers and ask whether they provide decisive grounds to reject MCI theory. I label the first objection the "blurry Animal/Person objection" and the second the "objection from Intuitive Dualism".

The blurry Animal/Person objection

Guthrie (2017) points out that an underlying assumption in MCI theory is that distinguishing human persons from non-human animals is intuitive and universal. This indicates that the concept of an animal that talks is MCI, since it involves a transfer between these domains (Guthrie, 2017, 75). However, Guthrie argues, the central element in detection of agency is attribution of intentionality, and this is involved in encounters with both humans and non-human animals. Moreover, whilst distinguishing the animate from the non-animate occurs automatically and is universal, discriminating between specific kinds of animated entities (i.e. humans, or non-human animals) is something subjects must learn (Guthrie, 2017, 75). To illustrate, Guthrie (2017, 75) for example says that young children readily think of animals as persons and that they must learn to think otherwise. Based on this, I envisage the following objection: MCI theory gets religious concepts involving alleged transfers between the Animal and Person categories wrong, since there is no intuitive, universal distinction between them. This means for example that the concept of an animal that talks is not MCI, something MCI theory fails to capture. Let us now investigate whether this objection undermines MCI theory.

A brief rehearsal of some points made in the previous two sections concerning intuitive ontologies and the functioning of HADD can help us. Guthrie's concerns are to some extent warranted, since the intuitive domains for Person and Animal in fact overlap: both, for example, involve intuitions concerning goal-directed movement and species essentialism (Boyer, 2000). Moreover, TOM is involved in interpretations of both human behaviour and the behaviour of non-human animals, since both are intentional agents (Barrett, 2005, 202). This overlap indicates that the boundaries between the intuitive categories for Animal and Person are to some extent blurry. But there also are areas which do *not* overlap. According to H. C. Barrett, despite some overlap there are important differences between the intuitions generated by encounters with humans and non-human animals. These reflect the distinct kinds of recurrent interactions our ancestors had with each type of agent. In short, predators and prey are "asocial", they typically just want to kill or get away, whilst conspecifics also have other goals and allow a much broader spectrum of interactions in addition to the "asocial" ones (Barrett, 2005, 202-203). Given the distinct recurrent challenges that interactions with humans and non-human animals posed, we have evolved a "categorizer" by which we discriminate between these kinds

of agents. This is HADD (Barrett, 2004). Moreover, after thus discriminating, further intuitions related to the distinct type of agent we have encountered (conspecific or non-human animal) arise produced by TOM (Barrett, 2005, 204, cf. Barrett, 2004, 3-6). This shows that there are distinct universal categories for Animal and Person even though these to some extent overlap. This is more plausible than Guthrie's claim that children must learn to make that kind of distinction. These points have implications for the blurry Animal/Person objection. First, the objection is wrong when stating that *no* universal distinction between these categories obtains. Furthermore, MCI theory does not get *all* concepts that involve a transfer between them wrong. Nevertheless, there is a danger that the theory gets *some* such concepts wrong. Let us explore this further to appreciate why.

The concept of an animal with human psychology, i.e. the sort of psychology we attribute when the categorizer has identified the agent as a conspecific, is in many cases, but not necessarily, counterintuitive. Such a concept is counterintuitive and therefore MCI if it attributes human psychology to an organism the categorizer would clearly identify as non-human. There is then a transfer between the non-overlapping parts of the Animal/Person categories. MCI-theory therefore gets right concepts that describe agents that we clearly identify as non-human, but which nevertheless are granted the sort of rich psychology humans have. But that does not mean this theory would correctly capture all concepts that transfer human psychology to an animal. There are many different types of non-human animals, which resemble humans to varying degrees. Concepts of talking non-human animals with some amount of hominid-like or humanoid features, such that our categorizer would yield conflicting or unclear identifications, might in fact not be MCI. We can imagine both clear cases and grey-zone cases, and the conflicted intuitions our mind would produce in the latter might create ambiguity concerning the MCI-status of concepts describing them.

Consider first the concept of a spider who talks. Encounters with entities exhibiting features statistically related to the morphology of spiders trigger activation of a specific evolved template. This in turn can lead to a fear response (Barrett, 2005). Attributing the psychology involved in human language-based communication to an animal such as this, for which we have a specific perceptual template and very restricted and automatic behavioural responses, seems counterintuitive. The concept of a spider that talks is therefore MCI. It involves a clearly "inappropriate" transfer of psychology to Animal. But what about the concept of a speaking member of the recently discovered hominid species Homo Naledi? Is that MCI, since there also in this case is a transfer of psychology to the Animal domain? This is less clear although Homo Naledi is non-human (that is, it belongs to the same genus but not the same species as humans).

Curiously, members of Homo Naledi might have co-existed with early Homo Sapiens. Moreover, they combine archaic features with features resembling modern humans (Scerri et al. 2018). We can only imagine how early humans might have intuitively categorized Homo Naledi. Perhaps the categorizer would yield conflicting outputs. It is then not clear that the concept of a member of Homo Naledi who talks (like a human) is counterintuitive. This is why the concept of an animal who talks is not necessarily MCI and that MCI theory therefore might get some concepts of talking animals wrong. But this is not because there is no difference in what intuitions we have concerning human and non-human animals. And it is not because MCI scholars are confused concerning how our categorizer works. It is rather because it is not clear how our minds would intuitively categorize this individual. In light of this, it is not always clear whether an animal with human psychology is counterintuitive. Therefore, MCI theory might have to make some adjustments concerning specific purported MCI concepts involving animals with humanlike features, at least in cases where that animal has features that lead our categorizer to yield conflicted or unclear outputs. But as the spider-example indicates, other concepts of speaking animals are probably MCI, since they attribute human psychology to agents we would clearly not identify as conspecifics. In conclusion, the problem for MCI theory here, if there is any, seems to be a greater need for precision and care in the construction of specific MCI concepts. However, that is no objection to the truth of the theory. So, the "blurry Animal/Person objection" fails to undermine MCI theory.

"The objection from Intuitive Dualism"

This objection applies to purported MCI concepts which arguably are more central to religion than those discussed above, namely the concepts of bodiless or immaterial persons. Guthrie (2017) points out that on MCI theory, the concept of a bodiless person is counterintuitive. This means many concepts of ghosts, gods, and spirits are counterintuitive. But, Guthrie objects, the evidence does not support this claim: in particular, the case of *intuitive dualism* indicates that such concepts are intuitive (Guthrie, 2017). We are according to Paul Bloom intuitive dualists in the sense that our minds readily think about minds and bodies as different and separate. Entertaining the notion of a bodiless mind is therefore intuitive (Bloom, 2004, 2010). This means that religious concepts which depict bodiless agents are intuitive (Guthrie, 2017, 76). The objection then is this: MCI theory provides an incorrect account of a broad range of religious concepts involving bodiless or immaterial agents by depicting them as

counterintuitive. Let us now investigate whether this objection succeeds in undermining MCI theory.

As a start, it is worth noting that MCI theory agrees with Bloom and Guthrie that notions of bodiless agents to a large extent are intuitive. The difference is that the former adds the caveat that agents lacking body are *minimally* counterintuitive by violating naïve physics (in section 4 we saw how such concepts to a large extent build on the intuitions the mind produces). We are therefore not discussing whether typical religious concepts are *massively* counterintuitive or *completely* intuitive but whether they are *minimally* counterintuitive or completely intuitive. By adding the caveat about minimal violation of intuitions, MCI theory has an advantage by offering a helpful way to see how typical religious concepts are different from concepts that are completely mundane and ordinary.[44] On the intuitive dualism theory, for example, it seems that the concept of a person walking down the street and the concept of a bodiless person walking down the street are equally intuitive. But clearly, these concepts are different. Whilst it is not clear that intuitive dualism offers a convincing account of what this difference amounts to, MCI theory has a good grasp of what the difference involves, namely violation of intuitive ontology.[45] We have seen that the empirical evidence tells us that such violations are memorable.

This brings us to a second response. K. Mitch Hodge (2008) argues that we are not intuitive dualists the way Bloom has argued. Hodge points out that Bloom's intuitive dualism thesis (2004, 2010) not only describes us as being drawn towards distinguishing minds from bodies but as being intuitive *Cartesian substance dualists.* Bloom (2010) argues that because we have evolved distinct systems for dealing with material objects and with agents, we intuitively think that bodies and minds are completely distinct, something which makes us intuitive Cartesians. Hodge (2008) argues that this is not correct. On the Cartesian version of dualism, minds are pure mental substances without any spatial location. According to this theory, which Descartes argued at length to defend, when the body dies, the mind, which always has been distinct from the material, continues to exist as a pure, non-localized thinking substance. Moreover, minds and souls are the same (Hodge, 2008, 391). But this is not the way

[44] In the next chapter, section 5, we shall see examples of religious concepts that are massively counterintuitive. Such concepts are found in theology and includes for example the Christian doctrine of the Trinity.

[45] Curiously, Bloom (2004) states, following Boyer (2002), that religious concepts are in fact different from other concepts in virtue of being counterintuitive and to some extent at odds with common sense. He even states, hearkening back to MCI theory, that the immateriality of such objects as ghosts makes them "interesting" and "memorable" (Bloom, quoted in Hodge, 2008, 391). So, it is not clear that Bloom intended intuitive dualism to offer another view of religious concepts than the one found in MCI theory.

we intuitively think about minds.[46] Among the examples Hodge provides to illustrate why, he points to studies of how subjects intuitively tend to represent the dead. For example, the study of funerary rites indicates that showing great care and respect for dead bodies is widespread. Moreover, there is good evidence for the commonality of burying the dead with a range of artefacts and possessions. If subjects intuitively thought of the dead as pure Cartesian minds, these behavioural patterns become puzzling. As intuitive Cartesians, they would think that the minds of dead subjects, where their personalities are housed, never have been material or bodily in the first place. On the Cartesian view, the minds of both the living and the dead are by definition purely non-material and mental. So, if this is the intuitive default-view we take of minds, why care for dead bodies or bury the dead with artefacts? (Hodge, 2008, 392-393). Furthermore, the study of mythologies and religious texts (such as Gilgamesh and the Bible) shows that the dead are frequently depicted as embodied and localized agents who reveal themselves to the living physically, rather than, what Cartesians would argue, by the use of something like telepathy (Hodge, 2008, 393-394). In light of this, Bloom's thesis faces challenges which undermine the intuitive dualism objection.

Recall also our discussion in section 3 concerning how our minds detect agency. I argued (following H. C. Barrett) that our minds have evolved a system which responds to some very specific perceptual cues when forming beliefs about the presence of agents. We do not indiscriminately attribute agency in the absence of any such cues, as this would be a quite costly strategy. Rather, we have a moderate bias towards false positives. What do we intuitively think

[46] In a defence of a slightly different version of MCI theory from the one presented above, Barrett (2008a, 319) argues that the notion of a mind triggers intuitions about spatiality. We intuitively attribute a spatial location to intentional agents with minds. By contrast, entirely non-spatial and non-localized minds are counterintuitive. This runs counter to the notion that Cartesian minds, which by definition are non-spatial, are intuitive. Consider also a set of studies Forstmann and Burgmer (2015) carried out. These provide evidence for a tendency to think that the mental and the physical *to some extent* are independent. For example, in a study involving the notion of physically duplicating living entities, subjects did not believe a physically duplicated hamster was mentally identical to the original hamster. They thought more physical than mental characteristics were retained in the duplicate. This indicates that they thought of the physical and the mental as to some extent independent. However, the duplicated hamster was still thought to retain some mental characteristics. This means subjects did not think of the mental as *entirely* independent of the physical. Furthermore, Forstmann and Burgmer also found that subjects tended to think of mental properties as linked or attached to physical properties. In a thought-experiment involving a reassembly (instead of a duplication) of the molecules that made up the original entity, subjects to a greater extent thought that mental properties would be retained in the reassembled entity. In a discussion of these results, Forstmann and Burgmer argue that their findings support the notion that we are intuitive dualists but depart to some extent from Cartesian dualism by indicating that subjects do not think the mental is completely independent of the physical (2015, 232-233). When it comes to the relation between the mind and the soul, a study carried out by Richert and Harris (2008) found that subjects tended to think of these differently. For example, the majority of subjects thought that the mind is subject to change and that it does not survive death. By contrast, the soul was typically thought of as not subject to change and not annihilated at death. Moreover, the mind was conceived as involved in cognitive functioning whilst the soul was linked to spiritual issues. These findings support Hodge's (2008) claim that subjects do not intuitively favour the Cartesian view of the mind, which sees it as identical to the soul and entirely independent of physical phenomena.

of the agents or the agency we sometimes attribute on ambiguous or incomplete perceptual evidence? The notion that we then attribute a *non-embodied* and *non-localized* agency is not plausible. That our mind "bets" on the presence of an agent in the absence of clear perceptual evidence (employing a "better safe than sorry" heuristic) does not mean that it thinks of an agent without a body and a localization. That would not make evolutionary sense. The "bet" can plausibly be construed as the bet that there is an agent somewhere *here* (i.e. a localized agent), since it is the presence or absence of localized (but perhaps hidden) physical agents with teeth and claws or spears and arrows that affect our fitness, not the existence or non-existence of purely mental agents who literally are nowhere in space or time. This further supports the claim that the notion of a bodiless, non-localized agent is not intuitive but counterintuitive, further undermining the intuitive dualism objection to MCI theory. In conclusion then, the intuitive dualism objection fails to undermine MCI theory.

As we have seen in the current section then, there is strong empirical support for MCI theory, and central objections to the theory fail to show that it is wholesale problematic. Instead, what these objections seem to indicate is that MCI-studies could benefit from improved designs.

Finally, a remark about the implications of the outcome of the MCI debate for the present investigation. In chapter 7, I argue against the epistemic justification of theistic beliefs, and (as mentioned in the final part of section 2 in the current chapter) a central premise there is that subjects in different cultural contexts will believe in different supernatural agents. To be more specific, the *virtue* of something I call the "MCI disposition" requires a tendency to form true beliefs and to avoid falsehoods with respect to a set of propositions about "MCI agents". The content of this set is strongly related to the amount and variety of MCI concepts of past and present cultures. It is since I find MCI theory compelling, that I in chapter 7 cash out the idea of such agents as MCI agents, and describe the disposition by which beliefs in these agents arise as the "MCI disposition". But this does not mean the argument crucially hinges on the viability of MCI theory. I suspect that the disposition I have mind can be reconceptualized using also other theories in CSR. For example, MCI theory is not the only theory in the field that offers a framework for accounting for the great cross-cultural variety of religious ideas, as the case of adaptationist and hybrid theories (e.g. Johnson, 2016 and Norenzayan et al. 2016) illustrate. Moreover, the next section introduces "context-biases" and discusses the importance of cultural input in explaining why people believe in different things in different contexts. This further illustrates how the central premise I make use of in chapter 7 can be defended by work in CSR.

6. Why not believe in Mickey Mouse or Zeus? Will Gervais and Joseph Henrich on context-biases

On the basis of what we have seen in the previous sections, we have an explanation of the ubiquitous presence of cognitively attractive concepts in human culture. But we also want to know why specific versions of MCI concepts tend to congregate in specific social and cultural contexts. Furthermore, and this relates to the issue of belief-formation, people do not randomly believe in any cognitively attractive deity but tend to form beliefs in those supernatural agents that people in their social and cultural context believe in. Why is that the case? The present section turns to a recent proposal about the importance of cultural learning mechanisms to answer these questions. We start by looking at an alleged problem for MCI theory, the solution to which directs us towards that proposal.

Some scholars in CSR do not see MCI theory as providing a sufficient explanation of religion.[47] Their objection revolves around a purported problem for the theory in distinguishing religious from fictional concepts and in accounting for why belief involving the latter concepts is not widespread. Consider Mickey Mouse. This character possesses abilities and properties that violate ontological expectations. Mickey Mouse is an animal with human psychology. Our categorizer would not think it was a conspecific. Therefore, Mickey Mouse involves an inappropriate transfer of psychology to the domain of animal. Such a minor violation should make the concept attractive. According to the MCI theory, we could therefore expect such concepts to occur in religious belief-systems. Why is that not the case? Why do people not believe in Mickey Mouse but believe in God, Zeus, Thor, and so on? Scott Atran (2002, 13-15) argues that a focus only on biases in human cognition cannot provide an answer to this. Theories with such a focus fail to explain why people often are strongly motivated to carry out costly acts related to their religious beliefs. Will M. Gervais and Joseph Henrich (2010) criticize MCI theory on roughly similar grounds and argue that MCI theory is inadequate. An adequate explanation of religion must in their view not only focus on the structure of concepts and which general kinds are cognitively attractive. An adequate explanation of religion must also highlight the social and cultural contexts in which subjects interact with each other and thus acquire information from which they build religious concepts. Henrich and Gervais argue that to

[47] These scholars see something like MCI theory as necessary in an account of the contagiousness of religious concepts, but they deny that this theory provides a sufficient explanation. By contrast, other scholars, such as Stewart Guthrie, reject MCI theory.

understand how subjects do this, we must focus on biases in cultural learning (2010, 383-389). We shall shortly look at this proposal. Let us first take stock of the criticism and appreciate how Boyer and J. Barrett respond to it.

When it comes to the issue of obviously fictional concepts that are cognitively attractive, Boyer and J. Barrett recognize that non-belief in them in contexts where they are common needs an explanation. Why do people believe in God and in spirits but not in Mickey Mouse? According to Boyer (2000), Mickey Mouse concepts and theistic concepts are not cognitively identical because the latter tend to latch on to intuitions related to *social interaction and morality* in a way the former fail to do. Subjects represent gods but not Mickey Mouse as agents with full access to *strategic information* (Boyer, 2000, 2002). Strategic information is any piece of information that activates our evolved cognitive systems for social interaction. What exact information that does this depends on the situation (and our goals, motivations, and so on). Importantly, Boyer continues, human minds automatically try to keep track of how much strategic information other agents have. Because we intuitively assume that human agents have *limited access* to such information, we must use cognitive resources to keep track of how much others know. In contrast, theistic (and other religious) concepts are concepts of agents with *full* (unlimited) access to such information. This makes gods relevant and it makes them seem powerful and important to us. Mickey Mouse concepts are culturally successful as the result of cognitive biases, but do not involve the notion of full access to strategic information and a moral concern for our behaviour. Therefore, we are unlikely to see Mickey Mouse (and similar fictional agents) as powerful and morally concerned. This is a possible answer Boyer and Barrett will give to the question of why such concepts do not occur in religious belief-systems even though their counter-intuitiveness makes them cognitively attractive (Boyer, 2000, 202-209, 2002, 171-180, Barrett, 2008b, and see Purzycki et al. 2012 for experimental evidence that subjects intuitively represent deities as agents with access to strategic information).

Critics of MCI theory do not seem satisfied by this response. Let us consider a further alleged problem for the theory, which points us to the proposal about cultural learning mentioned above. Suppose we can account in the above-suggested way for why concepts of Mickey Mouse and gods differ in terms of cognitive attractiveness. The latter but not the former concepts tap into processes by which we track how much strategic information others have. Let us grant that cognitive biases account sufficiently for why the latter but not the former type of concept occurs in religious belief-systems. We still want to know why people are so selective when it comes to which deity they believe in, when so many (equally) cognitively attractive deities are on the cultural marked. Furthermore, we should notice that this selectivity follows a

pattern we all can readily observe: subjects in context C believe in god G, which is culturally prominent in C, subjects in C* believe in G*, which is prominent in G*, and so on. This suggests that something more than biases towards cognitively attractive agent-notions is involved when subjects form religious beliefs, and that MCI theory does not account for this. Gervais and Henrich label this problem for MCI theory "the Zeus problem" (Gervais and Henrich, 2010, 385). Gervais and Henrich (2010) and Gervais et al. (2011) argue in favour of a more comprehensive account that combines a focus on cognition with a focus on cultural learning.[48] Here, in brief, is their proposal.

An adequate theory of religion must make use of two sorts of biases related to belief-formation. The first is a *content bias.* This bias attracts us to certain kinds of concepts in virtue of their content and structure. An appeal to such a bias is central to MCI theory as described above. It is in virtue of having minimally counterintuitive content that a concept is likely to propagate (increase its frequency). We also must introduce *context-biases.* These guide the learning of cultural information, and they nudge us to pay special attention to beliefs that occur with high frequency or that successful individuals in one's social context seem to have (Gervais and Henrich, 2010, 385-386, Gervais et al., 2011). The underlying idea here, which H.C. Barrett describes is that we do not only have evolved cognitive systems for handling adaptive tasks like detecting agents, facilitating proximity to caregivers, or reading minds. We also have systems that are adapted to *culture*, the environment of socially transmitted, potentially fitness-relevant information. This environment is our own production, and we constantly shape, develop, and alter it. Within this environment we find cooking recipes, literature, clothing-trends, beliefs about dangerous areas, tool-making techniques, and ideas about deities, spirits, and demons. As H.C. Barrett further argues, the reason we have evolved mechanisms for cultural learning is that cultural information is fitness-relevant; some pieces of cultural information affect one's reproductive success positively (Barrett, 2015, 224-227, cf. Boyd and Richerson, 1985).

The mechanisms we have for cultural learning make use of distinct cues when inferring what pieces of information to pay special attention to. *Frequency* or commonness is one such cue. Ideas, beliefs, and behaviours that are highly frequent or common usually (but not always) affect fitness positively. The usage of this cue therefore is part of how cultural learning mechanisms operate, something which creates a *frequency-bias*, a proclivity to conform to what

[48] As we saw earlier, many scholars in CSR highlight the importance of cultural input in proximate accounts of belief-formation or the construction of religious representations. Consider for example Boyer (2002), Barrett and Lanman (2008), and Van Leeuwen and Van Elk (2018). The current proposal can therefore be seen as agreeing with these on the importance of cultural input to account for the formation of religious beliefs.

is common in our context. By using frequency as a cue, we can also obtain beliefs and ideas that have no fitness-value. It is sufficient for the underlying mechanism to evolve that frequency largely pays off (in terms of fitness) while allowing occasional mistakes (Barrett, 2015, 228-229).

We also use the *success* of other individuals as a cue to what sort of beliefs and behaviours we should adopt. Our cultural learning mechanisms identify material resources, high social status, societal influence, and many friends as signs of success. They then use an underlying heuristic: imitate the successful because this can make you successful too. By believing, dressing, or acting the way they do, we might also (this heuristic says) obtain resources, status, and friends. The use of these signals as a cue creates a *prestige bias*, a bias in favour of imitating the successful. This bias also serves us well (or at least often enough for the mechanisms that use it to evolve). However, as H. C. Barrett points out, usage of it can also (depending on who we intuitively identify as successful and therefore good to imitate) lead us to adopt beliefs or to act in ways that are mal-adaptive or fitness-neutral, such as performing a rain-dance, adopting the drinking-habits of rock stars, or forming religious beliefs. This can happen because our mind does not know which factors cause success. This is opaque. Cultural learning mechanisms therefore simply *infer* that factors associated with the successful somehow are causally related to their success (Barrett, 2015, 229-231, cf. Boyd and Richerson, 1985).

The latter point about *inference* being involved when we navigate our cultural environment is important for our discussion of religion. The two above biases occur as the result of (intuitive) inferences our minds make when navigating the cultural environment and finding out what pieces of information to pay special attention to. Inference also is involved when cultural learning-mechanisms operate on those parts of the cultural spectrum we refer to as religion. These mechanisms nudge subjects to form particular religious beliefs in their particular context, by making an inference from the beliefs of the majority or the beliefs of prestigious individuals in that context. Given that what others believe is opaque, cultural learning mechanisms therefore require that we attribute beliefs to others first. Gervais et al. (2011) describe a specific strategy by which this happens, which is to use *credibility-enhancing displays* (CREDs) as cues to figure out what other people believe. Gervais et al. (2011, 392) provide the example of a cultural learner who observes a subject eat a mushroom. This is a CRED, a behaviour that reliably signals the presence of the belief that the mushroom is edible. It is a stronger signal than a behaviour whereby the subject simply says that the mushroom is edible and asks if you want to try. Religion involves many CREDs. Saying in public that there is no other deity than Allah and that Muhammed is his prophet indicates that you are a Muslim, but saying this *and*

abiding by the rules of Islam is a more powerful signal of such belief. Imagine how powerful the effect can be when one sees others respect costly religious rules, such as avoiding alcohol, fasting during the Ramadan, or praying five times a day. In highly religious societies, subjects are exposed to a large number of CREDs and therefore likely to attribute the requisite religious beliefs to the majority. In such societies, subjects are also likely to see CREDs in prestigious individuals (such as the religious or political leadership). Then obviously, conformist and prestige biases can have powerful belief-inducing effects. By introducing context biases such as conformism and prestige and by highlighting the use of CREDs and combining these factors with the content-biases described by MCI theory, we obtain a more powerful account of religion (Gervais and Henrich, 2010, 385-386, Gervais et al. 2011, Kirkpatrick 2005, 314-315, and for a more general account of biases in cultural learning see Barrett, 2015, 224-242, Boyd and Richerson, 1985).

We here have an explanation of why people are so selective about which specific deity to believe in, and thus a solution to the Zeus problem. The introduction of biases in cultural learning allows us to emphasize the influence of culturally contingent input from our social surroundings. Our minds reconstruct cognitively similar yet culturally distinct versions of such concepts from socially transmitted information. The *same psychological processes* thus yield beliefs about God, Zeus, ghosts, ancestors, and demons when implemented in different contexts. We return to this issue in chapter 7, where we look at the epistemic implications this has. The present proposal also allows us to pay attention to another important aspect of religion, which we return to in chapter 4: religion is a group-phenomenon. Having similar religious beliefs (and customs, clothes, and rituals) is a powerful way to mark the identity of a group. Exhibiting behaviours that indicate that one has the right beliefs is, in a religious context, an important way to signal that one is a proper member of the group and therefore to be trusted. One then receives the benefits of group-membership. Conversely, if one fails to signal that one has the proper beliefs, then other subjects are likely to trust one less, since one now is not identified as a proper member of the group (Kirkpatrick, 2005, 258-266, see also Norenzayan, 2013).

The next section considers a theory of religion with a more proximate psychological focus. We shall encounter another adaptation said to influence religious thought and behavior, *attachment.*

7. God as a haven of safety: Lee Kirkpatrick and attachment theory of religion

Lee Kirkpatrick, a central advocate of attachment-theory of religion (ATR), assumes much of the theoretical picture described thus far, according to which various cognitive biases account for why the notions of gods as well as many other supernatural agents are culturally successful. Such notions are around in the first place as the result of our tendency to animate the world and to attribute human features to it. Through social interactions in distinct cultural contexts people provide each other with information on the basis of which they construct mental representations of gods (and many other similar agents). These representations can then occur in many distinct thought-processes, thus becoming relevant. Kirkpatrick argues that religious beliefs and behaviors are by-products of how functionally distinct parts of the evolved human mind carry out distinct adaptive tasks. These tasks include agency-detection, social exchange, coalition-related behavior and *attachment.* The central claim Kirkpatrick makes is that attachment is a distinct process that sheds light on important aspects of the psychology of monotheist, especially Christian belief (2005, 52-74, 101-102). His contribution to the cognitive and evolutionary study of religion is to argue that attachment should be part of the larger theoretical picture, which we now already are familiar with. Furthermore, he also shows how one can combine a focus on epidemiology and cultural evolution on the one hand and a focus on proximate psychological processes involved in religious belief formation on the other.

According to Kirkpatrick, many different systems that evolved by natural selection have contributed to giving rise to and influencing religion. A central reason religion is so pervasive is how well religious ideas "co-opt" the functioning of our mind, whose systems evolved as solutions to recurrent challenges that did not have anything to do with religion in the first place (Kirkpatrick, 2005, 336-337). Attachment, the cognitive system Kirkpatrick focuses on, is an adaptation we share with many nonhuman animals. Kirkpatrick thus agrees with Stewart Guthrie that religion at least in part is a by-product of the functioning of adaptations that are not unique for humans. We saw that Guthrie grounds religion in an evolved perceptual strategy that non-human animals also use, which is to exaggerate the amount of agency in the natural world to avoid costly false negatives. Kirkpatrick (2005) argues that some aspects of religion involve another feature we also share with non-human animals (even geese), namely that an evolved attachment-system governs how offspring and caregivers adjust their level of proximity to each other. ATR is thus an example of theories of religion that in the words of Guthrie, "root religion in a biological matrix shared by other animals" (Guthrie, 2002, 38). The theory also

relates to both Guthrie and Boyer in claiming (as mentioned earlier in the chapter) that religion involves perceived relationships with purported supernatural *persons*, with whom believers assume they can communicate in various ways.

However, ATR also differs from the theories we have explored thus far in focusing more directly on the psychology of religious beliefs and experiences rather than cultural evolution and epidemiology. In other words, it has a more proximate focus. This theory thus has the potential to shed light on issues that do not figure directly into the foregoing theories, and it achieves this without departing from the insights they provide. Furthermore, not only is ATR about the psychology of religious belief in general. The focus is on monotheistic, especially Christian belief. This is obviously of interest given the focus in chapter 3 on how theistic beliefs (i.e. beliefs about the Abrahamic God) arise.

In order to understand ATR, we must first know something about attachment-theory, on which it is built. John Bowlby developed attachment-theory as an attempt to update and develop a psychodynamic understanding of cognitive development in children. His focus was on evolved proximate mechanisms that govern behaviour. In developing his theory, Bowlby made contributions that prepared the ground for the later emergence of evolutionary psychology. For instance, he coined the term "Environment of Evolutionary Adaptedness" (EEA), which subsequent evolutionary psychologists have emphasized. Bowlby was particularly interested in the relevance of animal behaviour to understanding human behaviour and he made extensive studies of the work of ethologists Nico Tinbergen and Konrad Lorenz. Bowlby was interested in a phenomenon Lorenz called "imprinting". Lorenz observed that newly hatched geese would follow any movable object in their vicinity. Lorenz interpreted this as an innately specified behavioural pattern triggered in certain periods of an organism's life by particular kinds of perceptual input. Tinbergen argued that scholars should apply such insights from animal ethology to the study of human behaviour. Bowlby applied Tinbergen's idea of fixed behavioural patterns (such as following-behaviour in geese) to the study of the behaviour of human infants. The result is attachment-theory, which today is a major framework for research in developmental psychology (Mysterud, 2005, 36-37, 91-93).

The central notion in this theory is that of an attachment-system, an adaptation that provides the organism with protection and nourishment by facilitating proximity between caregiver and offspring (Kirkpatrick, 2005). A crucial reason this system evolved is that proximity between offspring and parents is crucial to the survival of offspring. A proximity-facilitating system better enables the genes shared by parents and offspring to propagate than alternative systems that do not facilitate proximity. The attachment-system facilitates proximity

by generating behaviours. In humans, relevant behaviours include crying, shouting, and attempts to make parents lift the baby up. This occurs when the attachment-figure becomes distant or something in the surroundings alarms the child. Scholars view this system like a thermostat that triggers behaviours in such specific conditions and turns off again when the proximity-level regains an optimal level (Kirkpatrick, 2005, 25-30).

Attachment-behaviours are part of a cycle whereby the child explores its surroundings and subsequently returns to the attachment-figure. For the attached person, the attachment-figure is *a secure base* from which to explore and *a safe haven* to which to return. Attachment also refers to a psychological process by which an emotional bond develops between parent and offspring (Kirkpatrick, 2005). The bond that develops early in life eventually stabilizes and becomes an *attachment style.* An attachment-style consists of relatively stable beliefs about the self and expectations towards others. Scholars use the term "Internal Working Model" (IWM) for these thinking-patterns, and distinguish "secure", "avoidant, and "ambivalent" styles. The first involves an IWM characterized by positive attitudes towards the self and positive expectations towards the attachment-figure, who the subject expects to be loving and responsive. The second model involves negative attitudes towards the self and negative expectations towards the attachment-figure. Subjects in this category do not expect that the figure will respond with love and care. The third involves beliefs about the self as helpless and consists of conflicting images of the attachment-figure. The subject is unsure if the figure will respond. Importantly, though attachment-style arises as the result of early interactions with caregivers, it eventually develops into a general model that the subject intuitively applies in interactions with others. Working models that arise in childhood thus generalize and become entrenched thinking patterns (Kirkpatrick, 2005, 30-42).

We shall see how ATR builds on this theoretical framework when claiming that the influence of the attachment-system in some cases extends also to gods and other supernatural agents. First, a brief clarification. As Kirkpatrick points out, ATR does not accept the Freudian view of religion as an infantile neurosis that involves wishful thinking (Kirkpatrick, 2005, 19). Scott Atran seems to think so, because he interprets ATR as a "comfort-theory" of religion largely based on Freud. He also seems to assume that examples of terrifying deities such as those found in ancient Mayan religion undermine the theory (Atran, 2002, 72-78). This is not correct. ATR does not claim that all beliefs about deities involve attachment-dynamics. As we saw above, it argues in favour of including attachment in a larger cognitive and evolutionary framework where many other adaptations also are important. The theory is also consistent with

the central thesis in CSR that religion is the product of the *normal* functioning of human minds and therefore not, as Freud presumably thought, the product of pathology.

A central assertion ATR makes is that God[49] and other religious characters related to the monotheistic, Christian religion can become attachment-figures and that when this happens, then the functioning of the attachment-system affects how the subject thinks about and relates to God and these characters (Kirkpatrick, 2005, 52-55). This means there will be a behavioural cycle whereby God functions as a *secure base* from which to explore, and whereby factors that signal the presence of danger or that the level of proximity to God has been compromised, lead the attachment-system to trigger attempts to "return to God" as a *haven of safety* (Kirkpatrick, 2005). But what does it mean to be close to God according to this theory? Kirkpatrick argues that when God becomes an attachment-figure, we must understand proximity as a feeling of psychological closeness rather than actual physical closeness. The level of psychological proximity or felt closeness to God will then be monitored by the attachment-system. For example, the feeling that God has become psychologically remote might activate the system, which then influences the subject to engage in religious attachment-behaviour. A central example is prayer. If this does not work, the result can in some cases be thoughts about having been abandoned by God. This creates anxiety (Kirkpatrick, 2005, 52-68). Perhaps then it is not without reason that some Christians describe hell not as a place where demons torment them for eternity but as indefinite separation from God. Once one has formed a relationship to God, mere separation from him is thought of as intolerable.

Compared to an account based largely on agency-detection, the attachment-based notion of closeness to God provides a more psychologically realistic picture of how theists might feel something they take as divine presence. Even though, as we shall see in the next chapter, beliefs about the presence of God *can* in some cases arise from an agency-detecting process, theists normally feel that God is present in a subtler way. Based on the theory Kirkpatrick presents, we can envisage that God seems present to a believer in the sense that the believer experiences the sort of security an optimal level of psychological proximity brings about. This means it is not necessary to have one's agency-detector signal the presence of God. We shall also see that in some Christian contexts (described by Tanya Luhrmann, 2012), subjects also learn to employ specific methods to create and sustain a strong feeling of proximity to God (or Jesus).

[49] Not necessarily "God" the way the present investigation uses that term. However, given the many references Kirkpatrick (2005) gives to Christian belief, it seems that by "God" he primarily means the Christian God. For simplicity, we shall in the current section follow Kirkpatrick's usage of "God" with a capital G.

When it comes to how the attachment-system influences the formation of theistic beliefs, ATR proposes two distinct types of process. The first is a gradual process whereby subjects form the same religious beliefs as their parents, to whom they are securely attached. The second is about how subjects who were insecurely attached to non-religious parents convert, sometimes suddenly, to Christian belief in adult age. The first is called correspondence and the second is called compensation (Kirkpatrick, 2005, 134). Let us briefly consider them.

According to the correspondence-hypothesis, differences in attachment style formed through early interactions with primary theistic caregivers lead to corresponding differences in the way subjects think about and relate to God. This is because these interactions have given rise to a general attachment-style, which also in extension applies to God. As we saw above, attachment-styles involve internal working models (IWM) that consist of expectations and beliefs about the extent to which the attachment-figure will be responsive (to one's needs). This now also applies to God. Those with secure attachment to primary caregivers therefore tend to think of God as reliably responsive and caring. In contrast, those with ambivalent or avoidant attachment-styles do not expect God to respond to their needs (Kirkpatrick, 2005, 101-126).

Compensation is about how subjects with insecure attachment to non-theistic parents convert to theism in adult age. These subjects failed to experience well-functioning relationships with primary caregivers in childhood. This leads to the prediction that such subjects will seek surrogate attachment figures that can meet their emotional needs. A cultural and social context in which the notion of God is around makes it possible for people to turn to God as a possible surrogate (Kirkpatrick, 2005). This usually happens when subjects are introduced to a friendly religious environment that exposes them to the notion of God, who there is described in terms that indicate intuitively that he will make up a good surrogate (Kirkpatrick, 2005, Hood et al. 2009, 225-233). For example, the notion that "God/Jesus loves you just the way you are", might to these subjects come across as a powerful message. In these contexts, subjects in the insecure category are especially likely to become theists through what is known as *sudden conversion*. Kirkpatrick points to numerous studies that in his view powerfully reveal a link between insecure attachment to non-theistic parents and sudden adult conversion to Christian belief. He describes this as an emotion-based path to religion and points out how subjects in this category often use love-language to describe their relationship to God or Jesus. Becoming a theist is, on this theory, akin to falling in love with God (or Jesus) and of obtaining the feeling of being loved by God (or Jesus) (Kirkpatrick, 2005, 127-159, cf. Luhrmann, 2012, 101-131).

Conclusion

The evolutionary process has not equipped us with an adaptation for becoming religious but instead given rise to various cognitive systems that solve problems related to dealing with conspecifics as well as predators and prey. Religion seems to be a by-product of how these systems work. The logic of agency-detection and our tendency to anthropomorphize the world provide a basis for the origin of concepts of supernatural agents in human natural history. The cognitive biases we have investigated in this chapter, especially those dealing with social interactions, contribute to explaining why we are so susceptible to such concepts and help account for their cultural propagation over time. Our minds construct religious concepts based on (a) naive ontological theories that are tacitly assumed and (b) socially transmitted, culturally propagated information. The most cognitively attractive structure involves slight violations of naïve ontological theories. For example, the notion of God as an invisible person violates naïve physics, and the notion of Jesus rising from the dead violates naïve biology. We have adaptations for culture, that is, for living in environments whereby large amounts of information is socially transmitted. These adaptations give rise to specific biases that help account for the distribution of religious concepts: Conformist and prestige biases and the use of credibility-enhancing displays as cues to infer belief nudge subjects to form the religious beliefs characteristic for their particular cultural context and to ignore beliefs particular to other contexts. Finally, in Christian contexts, evidence indicates that God might function as a sort of attachment-figure to which theists form lasting bonds. With this general picture in mind, we are now ready to investigate how subjects become theists and how theistic beliefs arise.

Chapter 3. The Ontogeny of a Theistic Disposition and the Formation of Theistic Beliefs

1. Introduction and chapter-outline

The theories we have considered in the previous chapter provide an account of the human susceptibility to construct, use, and transmit a range of cognitively attractive concepts of purported supernatural agents such as ghosts, spirits, fairies, and deities, including God. We now move from a general focus on this susceptibility to the issue of how people become theists, that is, subjects who believe in God, the culturally specific deity of Abrahamic theism. Is theistic belief-formation among the phenomena CSR in principle can explain or actually explains given its current state? As mentioned in chapter 1 section 5, Aku Visala, Pascal Boyer, Helen De Cruz, and Johan De Smedt seem skeptical. The former two object to the notion of using CSR-research to explain individual, religious belief-forming episodes, and the two latter point to an alleged challenge for attempts to specify general belief-forming processes. This skepticism, if warranted, has positive implications for the externalist argument for neutrality, E-ENT. According to this argument, facts about belief formation are epistemically relevant, but CSR does not provide relevant facts about the formation of theistic beliefs. Research in this field therefore (since it also fails to shed light on internal factors) lacks epistemic relevance for such beliefs. On a strong reading, P3, the central premise in E-ENT, states that it is in principle impossible for CSR to contribute to explain how theistic beliefs arise in individual cases (statement 1). On a weaker reading, the premise states that research in this area does not currently contribute to explaining individual belief-forming episodes (statement 2). P3 also denies that CSR in principle can explain belief-forming processes (statement 3) and that it currently explains such processes (statement 4). We shall now argue against these statements, starting with the strong readings.

Here is a chapter-outline. We start in section 2 by looking at some methodological issues Visala (2011) points to concerning the distinction between ultimate and proximate explanations and what Visala calls “personal explanation”. I argue that applying general insights to

individual cases is more likely to yield explanatory results than to view each individual case as an entirely idiosyncratic event which requires a unique causal explanation. Moreover, what Visala (2011) calls "personal explanation" is unlikely to succeed because subjects are often wrong about why they believe certain things. Instead, CSR provides our best shot, since it provides general knowledge which together with information about the particular social and cultural setting in question can be applied to individual cases. In light of this, we have grounds to deny the strong readings of P3 in E-ENT and conclude that CSR in principle can contribute to explain how individual theistic beliefs arise and that it in principle can offer knowledge of general processes.

Section 3 attempts to show that CSR, if true, contributes to explaining how individual subjects become theists by offering general insights about this. I here introduce a distinction between how *people become theists* and how *theistic beliefs arise.* This distinction relates to work by Boyer (2002), Lanman and Barrett (2008), and Van Leuwen and Van Elk (2018), which distinguishes general religious beliefs from situation-specific religious beliefs and underscores the importance of the latter. I argue that to be a theist in typical cases is not merely to believe *that God exists*, but it is *to be prone to form specific theistic beliefs* in a range of situations. I call this propensity "the theistic disposition". This is my construal of what it means, in typical cases, to be a theist. The section offers an ontogeny for this trait, thus illustrating how CSR explains how people in general become theists. By "ontogeny", I mean an account of how this disposition gradually develops into a stable trait.

Section 4 discusses how we can envisage the theistic disposition as a dynamic trait whereby the intensity of the belief-production about God varies over time. The section thus sheds light on how we can see those who have the theistic disposition as theists. We then in section 5 turn to the theistic beliefs these subjects are prone to form. I describe three types of theistic belief: "intuitive theistic beliefs", "explanatory beliefs", and "folk-theological beliefs". We here see what kinds of beliefs subjects with the theistic disposition are prone to form and what kind of proximate factors that typically trigger their formation. The two former types of beliefs are situation-specific and largely based on intuitive thought. The latter beliefs are based on socially and culturally specific teachings to which subjects are frequently exposed in their context. In relation to these beliefs, we shall discuss the issue of "theological correctness" and whether theists believe in the official tenets of their religion. I argue that they do. Given frequent exposure to theology, subjects who are already prone to form intuitive and explanatory theistic belief are likely also to form folk-theological beliefs. By "folk-theological beliefs", I mean

beliefs in distorted or simplified versions of complicated religious teachings or correct versions of simple teachings.

The sections offer general accounts of how subjects become theists and how theistic beliefs arise. Given the combination of specific observations with this general knowledge, individual cases can be explained. In light of what the chapter argues, we can reject the strong and weak readings of P3 in E-ENT and endorse what I call "the explanatory adequacy thesis":

> CSR tells us how people become theists and how theistic beliefs form.[50]

2. Can cognitive science explain how theistic beliefs arise?

Aku Visala, appealing to Ernst Mayr, describes ultimate explanations as evolutionary accounts of gradual large-scale changes in populations. In contrast, proximate explanations focus on the development of traits in individuals. Visala applies this distinction to CSR-scholarship and contends that we largely find ultimate explanations there. CSR-scholars try to explain observable patterns in the propagation of concepts in populations over time. The explanations they propose appeal to evolved cognitive tendencies that constrain and influence a cultural evolution-process by which some concepts rather than others prevail. This enterprise focuses on cognition in order to provide an ultimate explanation of religion. It does not provide proximate explanations, that is, explanations of the development of religious beliefs or religious behaviours in individuals (Visala, 2011, 125-128). According to Visala, this has the following implications:

> "In explaining particular instances of religious behaviours and beliefs, ultimate explanations of cultural evolution are basically useless: if we want to explain, for instance, why is it that John believes in God, then we quickly realize that an answer in terms of cognitive similarities and cultural evolution is far from being sufficient. Indeed it might even be misleading (Visala, 2011, 127)."

[50] Further support for this conclusion is offered in chapter 7, section 2, where I, using a specific epistemological framework to be presented in chapter 6, specify a process by which theistic beliefs arise.

I agree that we cannot plausibly base explanations of how individual subjects form theistic beliefs solely on an ultimate account of religion. Such explanations must also deal with proximate issues. However, a closer look at the notion of a proximate explanation shows that CSR sheds light on such issues. We can envisage two kinds of proximate explanations, and only one of them resembles those Visala describes. The other is within the scope of what CSR-scholarship in principle can provide with respect to religious beliefs. Furthermore, such proximate explanations are applicable to individual cases even though they are general.

The distinction between ultimate and proximate is not primarily one between explanations of population-phenomena and individual phenomena. It is a distinction between explanations in terms of evolutionary functions and phylogenetic origins on the one hand and mechanisms and processes on the other (Mysterud, 2005, 30). Ultimate explanations answer why-questions and can tell us about the ultimate origin and function of a trait. Proximate questions answer how-questions and can tell us about the development of the trait in individual organisms during their lifespan. Nico Tinbergen (1963) uses this distinction when describing 4 types of questions an evolutionary science can ask about a behavior or a trait. The two first are proximate questions and the two latter are ultimate. The first asks how a behavior or trait develops over time in individuals. For example, how does the ability and propensity in birds to sing during spring develop? How do the mechanisms that produce following-behavior in ducks develop? These questions ask what the *ontogeny* of birdsong or following-behavior is. The second question is about causes or mechanisms. It looks at what factors cause a behavior to occur. For example, physiological events or environmental factors can play this role as releasers or triggers that produce bird-song or following-behavior in concrete situations. The third question is about what evolutionary function the behavior or trait has. In the case of the bird, this is perhaps to attract mates. The fourth question asks why the behavior or trait evolved. This is a question about *phylogenetic* origin. An answer to this kind of question tells a story about evolution by natural selection (Mysterud, 2005, 31-33).

The two first questions are proximate even though they are about something general. They might for example be about how a propensity to sing arises in birds. They are not about the individual life story of each bird. It is in virtue of focusing on the question *how* that these questions are proximate, not in virtue of aiming at unique cases. But that does not mean they cannot shed light on the latter. Correct answers to general how-questions are in principle able to shed light on individual cases because we can plausibly view individual cases as manifesting the operation of general processes and mechanisms, which are exactly what such questions aim at. But that does not mean knowledge of these processes and mechanism provides a sufficient

explanation of what occurs in an individual case. It *contributes* and we also need to introduce observations that are specific for the case we want to explain. This seems necessary to correctly subsume that individual case as a token of the general process or mechanism.

For illustration, consider the question why arctic hares are white during winter (Mysterud, 2005, 30-31 provides this example). An account in terms of the adaptiveness of the camouflage that white fur-colour provides in snow is ultimate and answers the third question mentioned above. An account of how that trait increased its frequency in a population of hares through natural selection in arctic environments is ultimate and answers the fourth question. These two accounts are insufficient as explanations of the whiteness of hares and not directly applicable to each individual hare. A sufficient explanation should also provide information about proximate processes and mechanisms that produce fur-change. An explanation in terms of a physiological process by which the pigmentation in the fur changes as the result of various factors that trigger this change is proximate. The same is true of an explanation in terms of a disposition in hares to change fur-colour seasonally. These explanations are answers to the first two questions described above. When they are in place, we know something about the nuts and bolts of fur-change in hares which we can apply to individual hares. Suppose we observe a given arctic hare, take notice its whiteness and decide that we want to explain this particular trait. Our combined ultimate and proximate knowledge, applied to this individual case, provides a good understanding of why it is white. This is because we lack grounds to suppose that the whiteness of this individual hare is the result of an idiosyncratic and unrepeatable process which requires a unique explanation, but we have strong reasons to believe that this trait is the outcome of general processes which also account for the whiteness of other hares in similar environments. Thus, even though proximate explanations answer general how-questions, they help shed important light on individual cases.

These points tell us something important about the prospects of explaining individual religious beliefs. The anthropologist Tanya Luhrmann, whose work we shall encounter again later in the chapter, sets the tone:

> "Despite the deep idiosyncrasies of personality and life path, when people feel and sense the divine, they do so in ways that can be detailed like a naturalist's observation on the flight of birds" (Luhrmann, 2012, 223).

Consider the question why S is a theist. The sort of scholarship we find in CSR does not provide what we can call a *super-proximate* explanation of this trait in this individual. By a super-

proximate explanation, I mean a detailed account of the life-story of a person, including idiosyncratic details about that person. That seems related to what Visala means with the term "proximate". Biological science does not provide such explanations of individual birds or hares either, but that does not hinder its ability to help explain traits that occur in individuals. We saw that it is possible to apply general proximate knowledge of the process by which fur-colour changes in hares to an individual hare we have encountered. This combination of general knowledge and observation relevant to a given case allows us to explain why *that hare* is white. Similarly, it seems in principle possible to apply general proximate knowledge of processes by which people become theists to an individual theist we observe. We can account for *why that person is a theist* if we plausibly can see what occurs in this individual case as manifesting general processes and mechanisms that produce similar results in other cases. Unless we are considering a pathological case, we lack grounds to deny the assumption that an individual case manifests general processes and mechanisms. If CSR can provide such general proximate knowledge, then such research can also contribute to explain a given individual case and thus help provide the answer to the question why S is a theist.

Can CSR provide such general knowledge? Scholars surely attempt to do this. This research does not focus only on the ultimate kinds of questions described above but also on the many how-questions that a study of religion brings up. As should be clear from our discussion in the previous chapter, this research surely is an attempt to say something about the nuts and bolts of religion. Before briefly rehearsing some of these, let us appreciate that central scholars in the area are also quite explicit in their criticism of what they see as a neglect of proximate questions in early examples of sociobiological and gene/culture co-evolutionary theories of culture (including religion). For example, Scott Atran (2002) and Pascal Boyer (1994) criticize attempts to provide an ultimate explanation of religion in terms of how religious beliefs, ideas, and behaviours affect fitness. Atran for example objects that such accounts are "mind-blind" in virtue of neglecting the mediating cognitive mechanisms and processes that enable religious ideas and beliefs to occur (Atran, 2002, 199-235). Boyer (1994) objects to theories of the transmission of ideas that focus solely on how fitness-relevant factors constrain such a process. He argues that we also need to deal with the ways in which our *cognition* constrains the spread of ideas. Many theories that view the spread of ideas in terms of cultural evolution do not, Boyer argues, have a psychologically realistic conception of the nature of ideas and the mechanisms by which they spread. Boyer for example objects to the notion that they, analogously to genes, spread by being replicated (Boyer, 1994, 14-16, 263-287). So, scholars in CSR see attempts to

provide ultimate explanations of religion without telling a proximate story about causal psychological mechanisms and processes as problematic.

CSR-scholars provide many accounts of cognitive processes, mechanisms, and faculties they allege to be involved in religious belief-formation and behavior, some of which we dealt with at length in the previous chapter. Scholars for example study agency-detection (Maij, Van Schie, and Van Elk, 2019, Van Leuwen and Van Elk, 2018), the memorability of MCI concepts (for example, Barrett and Nyhof, 2001, see also the review in Ch. 2, section 5), how religious activities affect brain-activity (Schjoedt, 2009 provides a review), and conceptual development in children (Kelemen, 2004, Barrett and Richert, 2003). The focus in CSR on proximate issues is also evident from the entries in the introduction to this field Slone and McCorcle (2019) provide, which presents empirical studies that have been influential. These examples of empirical work in CSR purport to shed light on distinct nuts and bolts of religion, a worthwhile endeavour in its own right, which also contributes to the larger goal of providing an ultimate explanation of religion. Some scholars in the field, such as Boyer (2002), Guthrie (1995), Atran (2002), and Kirkpatrick (2005), integrate work on such proximate issues in their ambitious attempts to provide an ultimate explanation of religion. This is what Visala sees as central to CSR-scholarship. We saw above that in his view, this research is largely an attempt to provide an ultimate explanation of culture in terms of underlying cognitive constraints.[51] Importantly for our present purposes, both those who attempt to provide an ultimate explanation and those who focus mostly on more proximate mechanisms and processes potentially provide proximate knowledge of religion. They both potentially answer general how-questions. This is important for our present purposes because, as I argued above, such general knowledge can tell us something about processes, faculties, and mechanisms involved in the formation of religious beliefs in individual cases. In the case of the arctic hares, we need to apply general insights about pigmentation-change to explain how a given hare became white. In the case of theists, ongoing research on the above-mentioned and further topics provide general insights about

[51] Visala (2011) uses "CSR" in a more restrictive sense than the current investigation. Visala's discussion focuses primarily on the work of J. Barrett, S. Atran, P. Boyer and D. Sperber, which he describes as the "standard model". The standard model largely consists of the MCI thesis and HADD (see Visala, 2011, 4-9). It is this model, it seems, which allegedly lacks tools to account for belief-formation in individual cases. In contrast, the present investigation uses the term "CSR" for a loosely delineated area in the study of religion that combines cognitive psychology, evolutionary theory, and anthropology in an attempt to explain religion as a natural phenomenon. CSR in the latter sense incorporates more than the former, given the many examples of attempts to shed light on nuts and bolts of religion. Notice, however, that something like the standard-model also seems capable of providing proximate knowledge. "CSR" in the sense of the standard-model is an attempt to provide an ultimate explanation of religion in terms of *underlying cognitive biases*, that is, in terms of biases that reflect the proximate functioning of the human mind. Thus, even though this attempt aims at the ultimate, it seems at least in principle able to provide important insights into proximate mechanisms and process as well.

cognition, which we can apply when explaining how a given subject became a theist. That does not mean that this research can explain fully and with certainty exactly how that subject became a theist. For example, we cannot deduce, a priori, based on general knowledge, what a specific individual will do or think in a given situation. However, such a level of precision seems hard to achieve in science in general and is therefore not something we should require in the first place. But given observations of the behaviour of an individual as well as some information about the social and cultural setting in question, we can obtain a plausible *likely* explanation of an individual case, where general CSR knowledge importantly contributes. That should suffice to satisfy our explanatory urges.

One might object that it is not only because CSR scholarship (or at least parts of it) attempts to provide an ultimate explanation of religion that this research does not explain belief-formation. Another reason is that this research makes generalized *statistical* claims about trends and tendencies in populations, and that it is a mistake to apply such claims to individual cases. Jon Elster points to two errors social scientists must avoid making. The first is to generalize from a single case (or just a few cases) to a population and the latter is to apply a statistical generalisation to a single case. For example, it is an error to infer from the presence of one aggressive man that all men are aggressive or to appeal to an alleged tendency in men to be aggressive in order to explain why a specific man was aggressive in a specific situation (Elster, 2007, 27). It could seem to be a similar mistake to think that a general tendency in humans to be religious accounts for why a specific person believes in God.

In response, CSR-scholars do not only make statistical claims about trends and tendencies. Instead, statistical claims are often starting-points for research in this area. For example, studies have revealed evidence that subjects in highly religious countries tend to distrust openly non-religious individuals (Norenzayan, 2013, 76-84). Though this finding is important, it is a mere starting-point for further research into the possible underlying processes that can account for such a tendency. This research can potentially provide proximate knowledge about how an inclination to distrust atheists arises in religious believers, and about the likely situational triggers that lead to discriminatory (or even violent) behaviour towards atheists. Norenzayan for example argues that an underlying cause of religious prejudice against atheists is that mechanisms for detecting free-riders (those who receive benefits through cooperation but without themselves contributing) identify atheists as free-riders. Why are atheists seen as free riders? Ara Norenzayan argues that this has to do with the intuitive belief that to be trustworthy one must believe that one is being monitored by a moralizing god (Norenzayan, 2013, 83). This example, which we discuss further in chapter 4, section 5,

illustrates a feature of research in social science described by Elster (2007). Elster points out that many endeavours in social science start with an observation of statistical relationships and then propose causal hypotheses in order to account for them. Such hypotheses should shed light on causal mechanisms and processes that account for the statistical observation. Elster further points out that we cannot plausibly apply general statistical claims directly to individual cases, but we *can* apply a plausible causal story to individual cases. We can explain what occurs in an individual episode by showing how it is a manifestation of a general causal pattern (Elster, 2007, 8, 15-37). Thus, we cannot explain why S distrusted S* (an open atheist) merely by pointing out that S is religious and that religious subjects tend to distrust non-religious subjects. But we can (as an example) argue that a likely explanation for why S distrusted S* involves the instantiation of an intuitive reasoning process in S, which produces the thought that one must believe in God to be trustworthy. Since it is possible for cognitive science to specify this sort of process and the situational factors that trigger it, it is possible for cognitive science to help account for why S distrusted S*. Then, we must make some observations about S and S* in the light of which we have grounds to believe that those situational factors are at play in this case, and we need to subsume what happens here as a token of a general process which tends to be triggered by those factors.

We might propose an explanation of this and similar individual belief-forming episodes in terms of the operation of a *psychological mechanism* thought to be involved in the relevant type of reasoning process. Bechtel and Wright (2009, 119) describe psychological mechanisms as functionally coherent composite systems that carry out a set of tasks or operations related to the regulation of behaviour or the processing of information. We shall in chapter 7 look at a related notion, supported by W. P. Alston (2005) and A. Goldman (2015), of psychological mechanisms as psychologically realized functional procedures. We will there see that an appeal to something like mechanisms in the sense Bechtel and Wright specify can help us specify a set of processes by which theistic beliefs arise. These processes involve functional procedures carried out by some of the evolved cognitive faculties CSR-scholars describe, such as the agency-detector and attachment. To return to our current example about distrust of atheists, this might involve a mechanism involved in regulating behaviour during social exchange. It might also involve something else. The point is that by appealing to purported knowledge of mechanisms, it is in principle possible to specify the processes they implement. In turn, this allows an application to an individual case thought to be a token of a given process. The puzzling statistical relationship (religious believers in highly religious societies tend not to trust non-religious subjects) is not our basis for such an explanation but a starting point for research

that reveals the underlying mechanism which together with relevant observations of the individual case in question provides that basis.

In light of what we have seen thus far, it seems that cognitive science in principle *can* shed light on theistic belief-formation in general by providing important proximate insights about general processes and mechanisms that are involved. We saw in the foregoing chapter that these processes and mechanisms involve naïve ontological intuitions on the one hand and exposure to socially transmitted information on the other. This indicates that statement 3 in P3 of E-ENT is wrong. It is not in principle impossible for CSR to explain how theistic beliefs form. We shall further undermine that statement in chapter 7, section 2, where I illustrate how three types of process that lead to theistic belief can be specified. I have also argued that the general knowledge of theistic belief-formation CSR is capable of providing *contributes* to explaining belief-formation in individual cases too. To be more specific, general knowledge about belief-formation together with information about the social and cultural setting relevant to the case in question allow us, at least in principle, to explain that case. This undermines another component of P3 in E-ENT, namely statement 1. Recall, this component of P3 states that it is in principle impossible for CSR to *contribute* to a sufficient explanation of how theistic beliefs form in each individual. This seems wrong in light of what we have seen. Based on the argument thus far then, we have grounds to reject the strong readings of P3 in E-ENT.

Does this conclusion concerning the explanatory capacity of CSR entail that we *must* appeal to cognitive science to explain religious belief-formation? Can we not explain belief-formation in other ways? Let us consider an alternative approach to this, which Aku Visala (2011) favours, called "personal explanation". Personal explanation assumes that subjects are rational and employs a belief-desire psychology which focuses on mental states like beliefs and reasons. An example of personal explanation of S's belief that p is that one simply asks S for S's reasons to believe that p and views those reasons as providing the explanation. Visala argues that the reasons subjects have often can explain beliefs and behaviours. It is therefore problematic to rule such factors out and systematically replace personal explanations with explanations in terms of automatic and non-conscious cognitive processes (Visala, 2011, 128-129).

In response to this, both personal and cognitive explanation in my view seem possible, but we have grounds to deny that personal explanation of religious beliefs usually provides our best shot at an answer. Instead, cognitive science usually provides our best shot. The reason personal explanation of religious belief is possible is that, as Timothy Wilson argues, even though this is not what usually occurs, subjects *can* in some cases find out the causes of their

beliefs through introspection. They can do this to the extent that the beliefs are formed in a conscious way (Wilson, 2002b, 106). That opens up for personal explanation whereby one accounts for the belief by introspectively tracing the relevant inferential steps. Moreover, we lack grounds to deny at least the possibility of forming religious beliefs in such a fashion. The reason cognitive science nevertheless provides our best shot at explaining religious beliefs is that these beliefs usually are not formed in that fashion. As for example Barrett and Lanman (2008) argue (see also Boyer, 2002, and Barrett, 2004), they are typically instead formed in ways strongly informed by intuitive processes subjects are *not* consciously aware of. Before returning to this point, let us first consider the prospects for personal explanation in more detail.

An initial challenge for personal explanation is that the subject might decline to provide any answer. Some theists object to answering questions about why they believe. They might even view their theist beliefs as brute doxastic facts that require no explanation. Furthermore, far from all theists would say that they have based their theistic beliefs on specific reasons or grounds. However, some theists point to reasons. These theists might say of specific arguments or considerations that these provide good reasons to believe and that these reasons are *their* reasons to believe in God. Even though this departs from the general trend in belief-formation Wilson (2002b), Barrett, and Lanman (2008) describe, it seems at least possible that in virtue of specifying reasons, the theists in fact explain their beliefs. This means that an approach whereby we see cognitive science as usually providing our best shot at an explanation does not rule out personal explanation. We should open up for personal explanation because reasons *can* cause beliefs and therefore are among the possible causal factors to consider. Reasons can cause beliefs in the sense that reasoning-processes where the subject considers reasons can lead to beliefs. With knowledge of such processes we therefore can explain beliefs (Lyons, 2009). This applies to beliefs that arise from conscious, slow, and deliberate reasoning-processes. It seems possible for subjects to use introspection to find out by what reasoning-process they formed such beliefs and then to reveal that process to us. Suppose the subject has formed a belief by carefully considering a set of reasons and then based a belief on them. The subject might then be able to trace the steps that terminated in this concluding belief. By asking why the subject believes in the conclusion, we obtain the right explanation. Thus, personal explanation is possible.

However, in many cases this kind of approach will not succeed because the processes by which beliefs arise are not like that. This is not only the case with respect to religious beliefs. Most beliefs seem to arise from fast, intuitive, and non-deliberate processes which subjects do not reliably tap into through introspection (Wilson, 2002b, Barrett, 2015). Wilson (2002b) and

Elster (2007) provide evidence for the view that we in general are not reliable sources of information about why we believe things or act in certain ways.[52] Many of our behaviours, thoughts, and judgements in regular every-day situations are the products of non-conscious automatic cognitive processes. We largely do not have introspective access to the processing that leads to these behavioural and doxastic results, and we fail accurately to trace the motivational or situational factors that influenced us. Nevertheless, we readily produce spontaneous explanations or stories that seem (to us) to account for these things. When asked about our beliefs or behaviour, we abstain from saying that we do not know (Elster, 2007, 126, Wilson, 2002b).[53] Rather, according to T. Wilson (2002b, cf. Barrett, 2015, 138), in such cases we *confabulate*. This means we invent reasons after the fact and falsely assume that these reasons are the real basis for our beliefs or the real causes of our behaviour. Confabulations can lead us to think that we are more coherent and rational than we in fact are. One might for example invent a false reason for why one chose a particular brand of jam at the supermarket, while the real cause was how advertisement primed the brain with that particular brand. But according to Wilson, this tendency to confabulate does not make us completely unreliable. When actions are the result of slow and careful deliberation or (as mentioned above) beliefs are the result of slow careful thinking, we might be able to explain them by bringing the responsible process to mind (Wilson, 2002b, 104-107). That is why personal explanation seems viable *in some cases*.

However, research in cognitive science provides strong reasons to believe that religious beliefs and behaviours largely involve processes that subjects *lack* the ability to trace accurately by introspection. In the accounts of belief-formation for example Boyer (2002, 341-379), Barrett (2004), and Barrett and Lanman (2008) provide, it is central that cognitively natural intuitions we usually do *not* have access to strongly constrain and influence the formation of religious beliefs. Barrett and Lanman (2008) for example contend that subjects first have intuitions and gut-reactions to events, and then, when asked about it, they form reasons. Furthermore, these reasons now seem to the subjects as explaining why the beliefs were formed:

[52] We readily also make mistakes when spontaneously explaining the behaviour of others. Simple belief-desire psychology often leads us astray. Evidence in social psychology for example suggests that we tend to make the "fundamental attribution error" when explaining the behaviour of others. We see their behaviour as the output of stable traits and dispositions and fail to appreciate how situational factors influence behaviour. This tendency has a self-serving function. For example, we might see a person who behaves aggressively as being inherently aggressive, but when we explain why we ourselves were aggressive, we focus on how situational factors influenced us. See Elster (2007, 186-188), Myers (2002, 84-85), and Buchanan and Powell (2018, 221-224).

[53] Elster (2007, 126) points out that the mind seeks meaning, looks for patterns, and constantly produces beliefs. He suggests that because of this, it is difficult not to form opinions about something when we ponder it.

> "We report our reasons as why we believe what we do, but the intuitions and emotional reaction might actually prove more reliable predictors" (Barrett and Lanman, 2008, 112).

This strongly indicates that asking theists why they believe will in many cases not produce accurate results, and it fits well with how I characterized confabulations above. The theists who do not admit ignorance about the factors that led to their beliefs but who instead try to provide a coherent explanatory story probably confabulate. This confabulation is likely to be a religious story told from the perspective of a believer, a story whose content is strongly conditioned by the cultural context. In an evangelical Christian context, it could for example be a story about how one "met Jesus" or "felt the presence of the Holy Spirit" and therefore "was saved" (see Luhrmann, 2012). Personal explanation, which is likely to tap into such a religious and theologically conditioned perspective, is unlikely to succeed in revealing the real causal factors that produced belief. In conclusion therefore, though personal explanation could at least in principle provide explanations of why subjects believe in God, cognitive science usually provides our best shot at an explanation.

3. The ontogeny of a disposition to form theistic beliefs

We saw in chapter 2 that the evolution of our cognition has made us susceptible to become religious believers. For example, cognitive biases make the human mind susceptible to obtain and construct religious concepts, and cognitive systems such as HADD, TOM, and attachment, can, in the right contexts, contribute to the formation of religious beliefs. Another contributing factor is cultural learning. Cultural learning is not the sole explanation, since religious concepts consist in part of intuitions the mind itself produces (Boyer, 1994). But as for instance Gervais et al.'s (2011) thesis about context-biases indicates, cultural learning importantly contributes to an explanation of belief-formation. In light of the discussions throughout the foregoing chapter, as well as section 2 in the present chapter, the combination of general cognitive biases and exposure to socially and culturally transmitted input offers the beginning of a promising explanation, in general, of how people become religious believers. This issue can be related to the first proximate question Tinbergen describes (see section 2). We now continue the exploration of this in more detail. I purport to show that central contributions to CSR, if true, explain in general how people become theists and contribute to explaining how this happens in

individual cases. In other words, we now move from a discussion of what CSR in principle can achieve to the question of what this research, if true, actually explains. First, a caveat about the assumptions I make in order to show this.

Given the importance of cultural exposure to theistic cultural material in a proximate explanation of how people become theists, one might object that CSR does not fully or completely explain how people become theists unless it also can provide the right ultimate explanation of how that cultural material arose and propagated in the first place. Can this field achieve this? CSR manifested in the work of scholars like Stewart Guthrie, Scott Atran, Pascal Boyer, Justin Barrett, and Lee Kirkpatrick clearly aims at explaining, ultimately, how religion arose. Recall our discussions in chapter 2 of biases in agency-detection (section 3) and cultural epidemiology (sections 4 and 5). We saw that MCI theory provides a compelling epidemiological account of how concepts spread over time. I argued that this theory helps us see why concepts of minimally counterintuitive, intentional agents with whom we can communicate in human ways tend to propagate. We then saw (section 6) that to understand why particular variants of such concepts become frequent in particular contexts, there is also a need for introducing biases in cultural learning. We shall see below how such biases contribute to belief-formation in a cultural context where *theistic* beliefs are common.

There is also research in CSR that aims at the cultural evolution and development of the specific cultural material we are interested in presently, namely the religious concepts, narratives, and teachings of Judaism, Christianity, and Islam. For example, we discussed (in Ch. 2, section 2) Ara Norenzayan's (2013) theory of how "Big Gods" emerged and gradually became strongly successful during the Holocene through cultural group selection. The Abrahamic God is a prime example of a Big God. There are also many other attempts in the current cognitive science literature to account for how the Abrahamic religious traditions emerged and developed, such as Wright (2004), Pyysiäinen (2009), Teehan (2010), Levy (2014), Baumard et al. (2015), Nicholson (2016), Uro (2016), Czachesz (2017), and Boyer (2018). For example, Baumard et al. (2015) and Boyer (2018) offer an alternative to the Big Gods hypothesis. They argue that the notion of a moralizing, powerful deity and the concern for the soul and its salvation which we find in Abrahamic theism emerged as the result of changing political and economic developments during the so called "axial age", from 600 BCE to 300 BCE. As another example of work on the Abrahamic tradition, Pyysiäinen (2009) looks at how the initial, anthropomorphic concept of Yahweh gradually developed from that of a tribal deity to eventually becoming a hugely successful monotheistic concept, focusing particularly on the Christian version. Teehan (2010) focuses on the ethical thought of the

Abrahamic traditions, arguing that it reflects an in-group/out-group bias which is part of our evolved cognitive makeup (more on this bias follows below and again in Ch. 4, section 4). As a final example, Nicholson (2016) explores the historical development of Christian doctrine in a psychological perspective, applying insights from CSR as well as social psychology, including research on in-group/out-group bias. A central thesis is that the formation of theologically complex doctrines and ideas such as the Trinitarian conception of God is within the scope of naturalistic, psychological analysis. By defending this thesis, Nicholson challenges what he sees as a problematic tendency in CSR to limit the scope of analysis to cognitively attractive religious ideas (such as MCI concepts) and thus to exclude the formation of theological doctrine from the purview of psychology (Nicholson, 2016, 10-12, cf. De Cruz and De Smedt, 2015, who also bring cognitive science and theology together).

We shall engage with some of these examples of work in CSR throughout the thesis, but it is not our primary focus presently. The reason I illustrate it here is that it is relevant to an assumption I shall make in the following. Given the need to introduce cultural exposure to theistic ideas in order to proximately explain how people become theists, and given that the latter is a substantive task of great significance for the present thesis, I shall make the assumption that CSR has or at least is in the process of providing a correct ultimate explanation of religion that combines a general knowledge of cognitive biases, cultural epidemiology, and relevant historical scholarship, and thus tells us how the *Abrahamic traditions* and their particular ideas emerged and developed.[54] This allows me to introduce exposure to theistic ideas to the proximate account below without first offering an (ultimate) account beyond the one offered in the foregoing chapter (which is about religious ideas more generally).

Let us now explore how people become theists. I begin with the issue of how we can construe being a theist. The emphasis in CSR on the practical nature of religion and the importance of intuitive rather than reflective thought will guide us. Pascal Boyer (2002) for example strongly emphasizes this, arguing that the primary difference between those who believe and those who do not is that the former tend to form intuitive, situation-specific beliefs about supernatural agents. These beliefs are produced when evolved systems, such as those related to moral thought and social exchange, are exposed to notions of such agents (Boyer,

[54] Recall, the present investigation explores the philosophical implications of CSR given that central contributions to this field are *true.* Since the above-mentioned examples illustrate relevant scholarship in CSR today, I see no reason for exempting them from being covered by this assumption. It is worth mentioning that if it turns out that CSR fails to explain fully how the specific Abrahamic traditions arose and developed, then that does not mean it cannot *contribute* to explain, proximately, how people become theists. But it would mean that such research does not fully or completely explain this, since it fails to account for all the factors a proximate explanation requires.

2002, 358-361). According to Boyer, most of the beliefs subjects form about such agents are not about their existence or what properties they have (such as power or omniscience), but they are about issues that are important in a given situation. Moreover, such beliefs involve thoughts about the supernatural agent's relation to the believers. For example, the belief that a given ancestor did not approve of a given action here and now is more representative for the religious beliefs subjects typically form than the belief that "ancestors exist" (Boyer, 2002, 358-359). Justin Barrett and Jonathan Lanman (2008) also emphasize the practical nature of religious belief, arguing that the experiences of the presence of agency that HADD sometimes generates interacts with beliefs in the existence of a given MCI agent, the result being that specific beliefs about that agent arise (Barrett and Lanman, 2008, 116). Thus, the argument goes, HADD, or other purported evolved systems help bring about a large number of specific religious beliefs about MCI agents in various situations (Barrett and Lanman, 2008, 116). HADD is the same across cultural divides since it is part of our evolved makeup. But which specific MCI agents it helps produce beliefs about depends on the cultural context. Finally, consider what Neil Van Leuwen and Michiel Van Elk (2018) call "general religious beliefs" and "personal religious beliefs". A general religious belief is a belief in the existence of a given supernatural agent. An example would be, "god exists". By contrast, a personal religious belief relates a general religious notion (such as "god") to the believer personally, as in the belief that "god loves *me"* (Van Leeuwen and Van Elk, 2018, 1-2, 12-13). On the model Van Leeuwen and Van Elk defend, intuitions about agency that are triggered in specific situations contribute together with an already formed general religious belief in producing personal religious beliefs: from agency-intuitions, the subject can "transition from merely having the general belief *that God*[55] *exists* to having the personal belief that God visited *me*"{emphasis original} (Van Leuwen and Van Elk, 2018, 14).

These proposals importantly highlight the practical nature of religion. Applied to our concern with theistic belief, they indicate that theists typically tend to form beliefs *about God* instead of only believing *that there is a God*.[56] Moreover, these proposals underscore the social nature of religion. The beliefs about God that theists typically form are not about abstract issues. They are about what occurs in social situations, and they relate to the emotionally rich and often life-long relationships to God these subjects form (Kirkpatrick, 2005, Luhrmann, 2012). In

[55] Not necessarily what I mean by "God" (as specified in Ch. 1, section 2).
[56] Having only the belief that God exists is sufficient for being a theist. Therefore, a person who has this belief but never forms any further beliefs about God is a theist. However, my point here is that typically, theists do not only have this belief but tend to form situation-specific beliefs about God.

light of these points, being a theist typically involves having a tendency to form beliefs rather than having one particular belief. I therefore suggest the following general characterization of what being a theist typically involves: to be a theist is to have a belief-forming habit whereby one is prone to activate an intuitive theistic concept and to form intuitive beliefs about God in a wide range of concrete practical situations. Moreover, when such a propensity is in place in a subject, God seems real to that subject. This propensity thus undergirds something like a general belief in the existence of God. Let us call this belief-forming habit the *theistic disposition*. This then is how we shall construe *being a theist.*

How does this disposition arise? Is this a process that occurs prior to the formation of practical, situational theistic beliefs? In other words, do people first form the theistic disposition and then, when it is implemented in specific situations, have it trigger beliefs about God? We have grounds to deny this. In light of the argument of Boyer (2002) and Barrett and Lanman (2008), the process works the other way around: subjects first begin to form intuitive beliefs about God, and as a result of this they gradually become prone to form such beliefs. Furthermore, when this propensity is in place in the subject, then we can attribute a general belief in the existence of God. As Boyer (2002) for example argues, it is not correct that subjects in a small-scale society where the belief that spirits roam the forest is common first form the belief *that spirits exist* and then, and only then, form specific situational beliefs about the spirits while in the forest. Instead, frequent inferences about spirits in specific situations, produced by evolved systems parasitized by spirit-concepts, gradually make the general idea of spirits in the forest increasingly plausible. Specific inferences about supernatural entities thus come first and their occurrence over time gradually strengthens a general belief in such entities (Boyer, 2002, 361-363).[57] This will be a central guiding idea in the following account.

[57] As mentioned above, Van Leeuwen and Van Elk (2018) also strongly emphasize the practical nature of religion. They for example plausibly interpret the tendency to form "personal religious beliefs" as more central than having a "general religious belief". Notice that their explanation of how personal beliefs arise involves purported inferences in the subject *from* general beliefs, in combination with situational intuitions, *to* personal beliefs (Van Leeuwen and Van Elk, 2018, 14-18). But how do general beliefs arise in the first place? Van Leeuwen and Van Elk here offer a slightly different picture from the one sketched above. They argue that such beliefs arise through cultural learning: one obtains these, for example, by talking to others in one's cultural context (Van Leeuwen and Van Elk, 2018, 11-13). But as our discussion of MCI theory and the Mickey Mouse and Zeus problems (see Ch. 2, section 6, see also concerns raised by Kirkpatrick, 2005, 117-122) indicated, mere cultural exposure to religious ideas is not sufficient to explain how people form religious beliefs. This applies also to general beliefs. Though it is plausible that cultural learning is involved, pointing out its involvement does not explain how the general beliefs arise. We might then instead ask, "how do subjects form general religious beliefs through cultural learning"? We could continue, "what proximate processes are involving in enabling subjects to obtain and use information transmitted to them during cultural learning?" Below I shall indicate what role I believe cultural learning plays in accounting for how subjects become theists and thus form a general belief in the existence of God.

A central reason specific inferences are likely to be prior to general ideas has to do with psychological relevance (see Ch. 2). The general notion that God exists, independently of any thoughts about how God relates to our lives, does not appear to have much psychological relevance. That belief is therefore not likely to stand on its own without support from intuitions which can help sustain it (Barrett and Lanman, 2008). By contrast, thoughts about what God does, believes, or wants from *us* are highly relevant, given how we evolved to be quite sensitive to socially relevant information. The general belief therefore appears to be a by-product of the specific beliefs rather than the other way around. Moreover, it seems to become established in the mind gradually rather than abruptly. As Boyer argues, specific beliefs gradually make plausible the notion that the supernatural agent in question is real (Boyer, 2002, 358-364).[58]

Cultural learning plays an important part in this process by exposing the mind to socially transmitted information that contributes to the construction of the theistic concept. Subjects do not obtain a full-blown religious concept in this way, since parts of its content are provided by the intuitions the mind itself produces (Kirkpatrick, 2005, 122, Boyer, 1994, see also Ch. 2, section 4). Rather, through social interactions in a theistic cultural context, the subject acquires various pieces of information that lead the mind to construct the concept of God. When encoding this information, the mind automatically relates it to the ontological domain of persons. The socially transmitted information refers to various statements or ideas about this person, such as "Jesus is the son of God", "God created the universe", "Allah's prophet is Muhammad", or "Yahweh entered a pact with Israel". Such ideas personalize God. Importantly, it is the socially transmitted information that suggests properties that violate naïve ontological expectations, for example by describing God as being able to listen to all prayers, being invisible, or able to perform miracles. By contrast, thoughts about God having a mind or believing things are intuitive and do not need to be learned (Kirkpatrick, 2005, 123, Boyer, 2002). The result of encoding this information, all of which one might not fully remember or easily get right, is a complex cognitively attractive theistic representation. We can view this representation as a sort of "micro-theory", a set of assumptions, ideas, and beliefs that occur in an interconnected complex cognitive structure (Boyer, 1994, 67-71).

Subjects are typically exposed to a large number of religious ideas, only a few of which become part of their religious belief-structure. What singles out "God" as special in some minds and therefore accounts for why that concept successfully latches on to intuitive, practical

[58] This means that even though subjects could deduce and form a general belief, "this spirit exists", from a specific belief, "this spirit was here now", the former belief nevertheless is not established firmly in the mind unless the subjects have formed many examples of the latter belief.

thought, whilst other cognitively attractive concepts do not? In typical cases, early interactions with primary caregivers who themselves have religious belief are involved in guiding the subject to treat some concepts as special (Hood et al. 2009, 114, Kirkpatrick, 2005, 114-116, Boyer, 2002, 365).[59] In fact, the best predictor of what religious beliefs subjects entertain as adults is the religious beliefs of their primary caregivers (Kirkpatrick, 2005, 115). In light of this, *theistic parents* are the primary sources of the information that leads the subject not only to represent the concept of God but to have this concept as the central religious representation.

This representation will not only be treated as special by the child given that the primary caregivers are the sources of the ideas that lead the child to construct it. In addition, the content of this representation will to some extent be affected by the content of automatic thoughts the child has about the caregivers. As we have seen (Ch. 2, section 7) Kirkpatrick (2005) argues that an internal working model arises gradually from interactions with primary caregivers. This model develops into a thinking-pattern that guides how the subject navigates interactions with primary caregivers and eventually also subjects in general, including romantic partners and friends. Religious concepts such as that of God entangle themselves with these thinking-patterns, which the attachment-system produces. The result is that attachment-related intuitive thought and thoughts of God become entangled (Kirkpatrick, 2005, 102-103, see also Hood et al. 2009, 134-135). The working model consists of tacit thoughts about the self (such as whether one is worthy of being loved) and beliefs about others (such as whether they are reliable, caring, and responsive); studies have shown that these attachment-related thoughts affect what sort of God-concept subjects construct (Kirkpatrick, 2005, 110-111). Since the working-model arises from interactions with primary caregivers, the latter interactions can be said to influence the

[59] This applies to the most common category of believers, namely those who form religious belief in childhood and whose parents are religious believers. In Ch. 2, section 7, we saw that Kirkpatrick (2005) distinguishes this category from another one, which consists of subjects who convert to a given religion in adult age. It will be recalled that these subjects typically have non-religious parents and that subjects tend to have insecure attachment to them. Moreover, in contrast to those whose beliefs reflect the beliefs of their parents, subjects in this category, which makes up a small proportion of all religious believers, often become religious through what is referred to as "sudden conversion" (see Kirkpatrick, 2005, 129-130). An explanation of how the subjects in this category become theists therefore needs to assume social interactions with other subjects than primary caregivers as the central event which singles out "God" as a special concept. As Hood et al.'s review indicates, conversion-studies point to the importance for the potential convert of forming social relationships to a person who is part of a religious group. Such a person, who is often referred to as "the advocate", helps introduce the potential convert to the religious group and functions as a door-opener to the religious life there (Hood et al. 2009, 230-231). Moreover, the phenomenon of sudden conversion might have some particular psychological aspects that are not present in the formation of religious beliefs in the statistically most common cases. In particular, Kirkpatrick points us to the phenomenon that subjects use their belief in God to compensate for lack of good, stable, relations to significant others earlier in life. Those who convert to religion as adults tend to have had bad relationships early in life, and God to them will seem to be the opposite of early caregivers (see Kirkpatrick, 2005, 127-159). By contrast, the general and most common pattern is that secure attachment to religious parents contributes to bringing about a conception of God that corresponds with one's conception of one's parents (for example as being responsive and loving) (Kirkpatrick, 2005, 101-126). In our account, for simplicity, we focus on this category.

particular idiosyncratic conception of God a given subject will form. For example, children with controlling (theistic) parents might think of God as more controlling than children with less controlling theistic parents. Moreover, statistically they are more likely to lose the belief in God later in life. By contrast, children with secure attachment to theistic parents will also tend to become securely attached to God later in life, and these subjects are also more likely to remain theists throughout life (Kirkpatrick, 2005, 102-106). This tells us that interactions with primary caregivers influence the content of the concept of God and lays a foundation for how well the child eventually will relate to God. This helps us account for individual differences in how a culturally specific religious concept is represented in a given culturally established group.

This brings us to the larger social and cultural context. Particular aspects of the wider social and cultural surroundings, which the child eventually notices, also influence his or her representation of God, further indicating that it is special compared with many other cognitively attractive concepts that the child also constructs as the result of socialization into the social and cultural world. The child will for example eventually notice that not only his or her parents but also many individuals in the wider surroundings seem to treat the notion of God in a special way, exhibiting signs of commitment to God. Credibility-enhancing displays ("CREDs") are behavioural displays which reliably indicate that subjects have particular beliefs. These are behaviours one would likely not perform if one did not have the corresponding belief (Gervais et al. 2011, 392). By observing other subjects in the social and cultural setting frequently perform religious rituals involving the notion of God, the child will intuitively think that these subjects have the corresponding belief in God. Then, conformist and prestige-related biases produce a pressure to adopt similar beliefs. This reflects how adaptations to culture (Gervais et al. 2011, Barrett, 2015, Boyd and Silk, 2015, Ch. 2, section 6) come to expression in a theistic cultural environment.

Introducing these factors helps us see how the concept of God will tap into intuitive thoughts by which the subject forms a cultural identity and navigates inter-group interactions. The central religious representation the child entertains will be affected by intuitive thought by which the evolved human mind interprets group-membership. Importantly, it will be related to intuitive thoughts about the *in-group*. The thought then arises of being part of a group that *follows God*. The subject's dawning awareness that there are other people out there who do not follow God not only helps fortify his or her identity as member of the in-group, but it also will affect intuitions about the *out-group* (Nicholson, 2016). We shall look closer at this dynamic in chapter 4, sections 4 and 5, but given its relevance presently, some points are due. The central aspect of groups is that they are sets of individuals who represent themselves and each other as

having similar goals which they must pursue collectively (Boyer, 2018, 43). When individuals represent themselves in that fashion, they assume a boundary that separates those on the inside from those on the outside. This leads to a bias whereby one thinks and acts differently towards the in-group and the out-group. For example, one is more distrustful towards the out-group, and is liable to react with less empathy to the sight of their suffering (Teehan, 2010, Boyer, 2018, 38-41, Tomasello, 2016, 91-92). In light of this, we can see how the concept of God in the child latches on to the intuitive ways in which the boundary between in-group and out-group is created and maintained. This further contributes to singling it out as special.

We have seen how the concept of God is constructed and why the mind of the subject we imagine will single it out as special. Let us now explore how the relevance of this concept increases in the mind over time, thus gradually bringing about the theistic disposition. Here the practical inferential activity Boyer (2002) describes (discussed earlier) is central. As we saw in chapter 2, section 5, the notion of *psychological relevance* plays (together with memorability) an important role in making a concept cognitively attractive. We saw that relevance relates to the cognitive ease with which subjects can employ the concept, how readily it leads to new inferences (i.e. its "inferential potential"), and how many systems it can trigger (its "aggregate relevance"). Relevance is also a property that can increase as the result of such inferential activity. Barrett and Lanman (2008) for example argue that when a religious concept is activated, then the memory of the agent it depicts is strengthened. Specific events whereby thoughts of an MCI agent suddenly arise establish that concept more firmly in long-term memory, and such memories gradually become part of one's "non-reflective cognitive resources" in ways that make the subject likely to affirm the existence of this agent (Barrett and Lanman, 2008, 118, cf. Boyer 2002, 358-364). As the result of how the subject thinks intuitively about God in a wide range of circumstances where many distinct evolved systems are activated, a wide range of associative paths to this concept are formed, making it increasingly easy to be reminded of God and form further beliefs about God in the future. These points indicate that we can distinguish two ways in which the relevance of the God-concept is increased, one temporary and one long-term.

A given case of activation of the intuitive concept of God temporarily enhances its relevance in the subject's mind. Let us call such activation-events "relevance-enhancing" processes. The beliefs that arise from relevance-enhancing processes are the intuitive thoughts our minds would produce anyway but with a theistic layer or gloss. By contrast, frequent relevance-enhancing episodes over time make God relevant in a lasting sense. After subjects, over time, in many different situations and in ways involving many distinct systems temporarily

enhance the relevance of God, this concept stands out in the mind in a lasting sense. God then has become a central part of the internal mental life of the subject and the concept is thoroughly embedded in the mind, likely to be activated at any moment. This relates to what I mean by the "theistic disposition", which is a propensity to think of God in a wide range of situations.

To get a clear picture of the sort of process that temporarily enhances the relevance of the concept of God, and which, if frequently implemented, brings about the theistic disposition, let us consider the following example. It involves the following elements: an adaptive challenge, an evolved system specialized to handle that challenge, and intuitions that arise in that type of circumstance. To this, I add the theistic representation and the assumption that it has been singled out in the mind as special in the way discussed earlier.

Consider the adaptive challenge *social exchange* represents. Leda Cosmides and John Tooby describe social exchange as interactions where parties provide each other with benefits, with the tacit expectation of reciprocation. A sort of tacit contract governs these interactions. This contract says that if you provide a person with a benefit, then that person should provide something to you in return. Another way to put it is that those who receive a benefit must pay a cost (Cosmides and Tooby, 2005, 584). Cosmides and Tooby argue that our minds contain an adaptation for this type of challenge, which has been a recurrent issue for humans and pre-humans during large parts of the Pleistocene. They call this adaptation is the *cheater-detector*.[60] A cheater is an agent who receives a benefit without reciprocating, thereby violating the tacit "contract" that governs social exchange. The cheater-detector produces intuitions about whether another subject with whom one cooperates is likely to be a cheater. These signals allow subjects to regulate behavior in ways that hinder exploitation during social interaction. The output of the cheater-detector consists of signals about possible exploitation, and these connect with programs in the brain that regulate emotions such as anger and resentment (Cosmides and Tooby, 2005, 591). According to Pascal Boyer, these emotions involve *moral judgement* and such judgement feels or seems objectively true to the judging subject. Boyer further suggests that this felt objectivity involves the automatic assumption that neutral by-standers with access

[60] As evidence for this system, Cosmides and Tooby (2005) present research using a particular version of the Wason selection task. This task in general is used to test how well subjects detect violation of a logical rule of the form, "if p, then q". Average performance is very poor when the task is presented in a purely logical form. But when experimenters presented a logically identical version of the rule that makes it seem that the task is about social exchange (such as "if you borrow the car, then you must fill gasoline"), performance was much better. Cosmides and Tooby argue that in such a version of the task, the violation of the logical rule seems to indicate the presence of a *cheater*, i.e. someone who violates the tacit contract in social exchange. The increase in performance allegedly reflects the workings of the cheater-detector specialized for social exchange (see Cosmides and Tooby, 2005, especially pp. 594-598). Scott Atran (2002) is critical of this usage of the Wason-task to provide evidence for a specialized cheater-detector (Atran, 2002, 283). Critics of evolutionary psychology, such as Buller (2006) and Richardson (2007) are also critical. For a recent discussion of the Wason-task, see Mercier and Sperber (2017).

to *strategic information* (see chapter 2, section 5 for more on this term) agree with the moral judgement. One assumes that everybody who do not have stakes in the situation and who knows what happened (in our current example, that one party in the exchange cheated) will form the same moral judgement. Boyer then argues that such an assumption also in extension applies to supernatural intentional agents, like deities, who the subjects represent as interested parties with full access to strategic information (Boyer, 2002, 215-218).

This indicates something important about how the concept of God becomes relevant by hitchhiking on a prosses that targets an adaptive challenge unrelated to religion. The assumption that everybody agrees extends also to God in virtue of how the concept of God in this case becomes activated and entangled with the moral intuitions. The assumption that everyone with such information agrees with the judgement lays the foundation for the intuitive belief that *God* agrees with it. Here is an illustration. Consider S, whose mind contains the concept of God. One day S provides Z with a benefit and Z fails to reciprocate. This turn of events engenders anger and resentment in S. This resentment includes a tacit moral judgement that Z is bad. S then intuitively believes that all neutral agents with access to information about how Z violated the tacit "contract" agree that Z is bad. The theistic concept latches on to S's intuitions about Z's badness, and this makes God temporarily relevant. It triggers thoughts about God disapproving of Z. Here is a sketch of the tacit inferential steps that eventuate in an intuitive theistic belief in S:

a. S's cheater-detector signals that Z is a cheater and generates resentment towards Z.
b. Tacit assumption that neutral agents who know what happened agree with S.
c. Formation of the belief that God resents Z.

This is an example of how concepts of supernatural agents "parasitize" our moral intuitions. Boyer suggests that it is in virtue of easily parasitizing these and many other intuitions that such concepts tend to become salient and relevant. Moreover, this process involves a confusion. The subject will think of his or her own intuitions that they are the opinions of a deity (Boyer, 2002, 218). In the above example, the theistic belief S forms is simply the intuitions the cheater-module engendered, but with a theistic gloss. This example thus describes how confusion enables the subject to think of unreasoned thoughts and intuitions in a new way because of how a theistic concept parasitizes them. The result is an intuitive theistic belief. Formation of such a belief temporarily increases the relevance of God. Thus, we have an example of a relevance-enhancing process. By contrast, the frequent occurrence of such processes over time, involving

many distinct systems dealing with a broad range of adaptive problems gradually create lasting changes. In other words, in this way, God gradually becomes lastingly relevant in the mind. When this has occurred, then we can attribute the theistic disposition to the subject.[61]

We have seen that the first part of the ontogeny of the theistic disposition is that people in the subject's surroundings introduce the concept of God to the subject, typically in early childhood. Exposure to CREDs indicating belief in God helps single out that concept as special even though other attractive concepts also are around. We have then seen an example of a relevance-enhancing process. This process *temporarily* enhances the relevance of the God-concept in the mind. As the result of frequent occurrences of such processes in a broad range of situations, the theistic concept eventually achieves a *lasting relevance* in the mind. Frequent intuitive thinking whereby the notion of God has entered the flow of thought has then created lasting mental patterns. These patterns influence how readily the mind activates the concept of God, making the subject prone to do this in a broad range of situations. The subject now has the theistic disposition: He or she is prone to think regularly about God. As a result of this, God seems real. We can now refer to this subject as a theist. This concludes our ontogeny of the theistic disposition and thus also the account of how subjects in typical cases become theists.

In light of this account, it seems that CSR, if true or at least approximately true, presently explains, in general, how people become theists. This undermines statement 4 in P3 of E-ENT, which states that CSR does not presently explain this. The falsity of statement 4 indicates that CSR also helps explain individual cases. This is because general knowledge of belief-formation together with relevant observations regarding individual cases explains such cases. Since, as I have attempted to show, CSR provides the general knowledge, it contributes to explaining individual cases too. For example, a given individual case can, given that one makes relevant observations of that case, be related to the general process by being construed as a token of this process. What we have seen therefore also undermines statement 2 in P3 of E-ENT, which denies that CSR presently contributes to explaining individual cases.

Recall our distinction between becoming a theist and forming theistic beliefs. We have construed being a theist as having the theistic disposition and accounted for how this disposition arises, thus providing an answer to the first of the two proximate questions Tinbergen describes.

---

[61] Notice that a similar relevance-enhancing process during social exchange in another cultural context could make a subject form the belief that another deity disapproves of the cheater. By imagining that this and other processes temporarily enhance another religious concept, we can account for how other deities become lastingly relevant. To account for this, we must alter the culturally specific input we add to the underlying cognitive process. In this way, we can account for how subjects become prone to form beliefs in other deities than God. For more on this, see chapter 7.

This is the question of how a given trait arises in the subject. But there is also the issue of how theistic beliefs arise, or in terms of Tinbergen's second proximate question, how a given behavior is triggered in a specific situation. The illustration of a relevance-enhancing process above sheds light on this. It indicates how a belief suddenly arises in a given situation. However, theists also plausibly form many other theistic beliefs. What beliefs are these and how do they arise? To answer this, I shall in section 5 provide an account of how three types of theistic beliefs arise in subjects who have become prone to form such beliefs. In this way we indicate that CSR explains both *how people become theists* (which the present section dealt with) and *how theistic beliefs arise*, thus further undermining statements 2 and 4 and supporting the explanatory adequacy thesis (see section 1). But before we get to that, let us briefly explore the functioning of the theistic disposition.

4. The theistic disposition as belief-forming propensity

A history of intuitive thinking about God leads to the formation of a habit, which is a lasting feature of the subject. This is the *theistic disposition* (TD). It has the following features:

(TD) is a relatively stable feature of the subject such that (a) there is a propensity for a theistic concept to figure as part of cognitive processes in ways that induce theistic beliefs, and that (b) God seems real.

We can understand the notion of a propensity to activate the theistic concept in a statistical sense as describing the likelihood that an event leads to activation of the theistic concept in a belief-conducive way. This provides a way to quantify the notion of psychological relevance. The more relevant God is, the higher is this propensity. Furthermore, in a subject with the TD, *God is lastingly relevant*. We can understand this as stable high likelihood of activation of the concept over time. How high must this stable likelihood of activation be for us to plausibly think of the subject as a theist? The definition leaves this open because it is problematic to assume that the propensity must reach a specific threshold for an attribution of the TD to a subject to be correct. Rather, there are clear cases and grey-zone cases. We therefore cannot always straightforwardly divide subjects into theists and non-theists, as if this is a binary category. Furthermore, this means that in the process of becoming a theist, the mind of the

subject must pass through a grey zone of propensity-levels which do not straightforwardly allow an attribution of the TD.

To better appreciate that the distinction between theists and non-theists sometimes can be blurry and that becoming a theist often happens gradually rather than abruptly, let us imagine that we have a control-board with two knobs. The first controls propensity, i.e. how likely it is that the theistic concept figures as part of an intuitive cognitive process.[62] The second knob controls the extent to which God seems real. This relates to our discussion in the foregoing section, and the contention there that a general belief in God arises gradually as a by-product of a tendency to form specific beliefs. The idea that a knob with various levels controls how real God seems illustrates that the seeming reality of God can change in strength. We can liken this seeming reality to a light whose intensity can be altered using a dimmer. With a dimmer, turning the light on and off is a gradual affair. The above ontogeny of the theistic disposition indicates that becoming a theist is more like gradually turning up the intensity of the light than flipping the switch to "on".

Not only becoming a theist, but also being a theist will in many cases involve variations in intensity over time for both the propensity and the seeming reality. It is therefore not akin to simply having the light on but rather something like putting the dimmer in a dynamic mode where the intensity varies. This feature of the religious life relates to the behavioural cycle Kirkpatrick (2005) discusses, which we looked at in chapter 2, section 7. The feeling of psychological closeness to God varies throughout the life of the theist, who might seek to retain a feeling of optimal closeness through prayer when this feeling starts to wane and God feels remote. In subjects with secure attachment to God, these variations are not likely to alter how real God seems, since attachment-behaviour (such as prayer) usually helps subjects regain closeness. God then again is a secure base (Kirkpatrick, 2005, 61-70). But since these variations relate to variations in the intuitive beliefs formed about God, which the seeming reality of God builds on, they potentially can impact on how real God seems. This is especially likely in those with insecure attachment to God. Their patterns of intuitive thought about God is a token of a more general internal working model leading to trust-issues and fear of abandonment (Kirkpatrick, 2005, 33-36). Kirkpatrick points out that insecure attachment predicts forms of adult religiosity that are unstable and liable to variation compared with secure attachment.

[62] This knob does not control the likelihood that the subject will process a theistic concept using any kind of procedure. For example, an atheist might think often about God using a reasoned analytical procedure. That does not mean the theistic concept the atheist then uses is likely to latch on to intuitive and automatic cognition in belief-conducive ways, which is what the knob for relevance controls.

Moreover, insecure attachment is associated with increased risk of loss of religious belief in adult age (Kirkpatrick, 2005, 137). Let us explore this. How can changes in intuitive thought-processes alter how real God seems? Let us first look at changes that attenuate the seeming reality of God.

We imagine a subject with the TD whose feeling of psychological closeness to God has weakened. This experience is in the light of attachment-theory likely to induce attachment-behaviour in the form of attempts to return to God as a "haven of safety". We imagine further that the subject's religious exertions fail to offset the feeling of God being psychological absent. Then, the subject, still prone to think of God and familiarized with religious (culturally prescribed) imagery, might form the thought that God has abandoned him or her, and this belief might linger. The initial impact here is on the subject's assessment of his or her relationship to God, a deity who still seems to exist despite not seeming present to the subject. Something that appears real here appears remote. But this felt absence of God also has potential impact on whether God seems real, since, as we have seen earlier, this seeming reality is the product of intuitive, situational thoughts about God. If the latter thought, which includes attachment-governed thought, changes, then that potentially changes the seeming reality of God. This indicates that by appearing psychologically absent, something attachment-related processes can account for, God eventually appears less real to the subject. We can better appreciate how such changes can impact the seeming reality of God using the metaphor of the light-dimmer that turns down the intensity than by imagining a switch being flipped to "off".

What factors will accentuate the seeming reality of God? The variations in the feeling of closeness to God that theists experience are culturally distinct expressions of the underlying natural cycle attachment-relationships in general involve, whereby the absence of the attachment-figure is followed by attempts to come back again to this figure (Kirkpatrick, 2005). In subjects with secure attachment to God, a temporarily lost feeling of psychological closeness to God is likely to be followed by feelings of having returned to God. Prayer but also interactions with various religious artefacts like icons, relics, statues, and so on, remind these subjects of God, thus facilitating closeness (Kirkpatrick, 2005, 59-60). Furthermore, negative life events such as illness or disasters function as triggers that nudge theistic subjects to engage in attachment-behaviour, leading to potentially intense feelings of closeness to God (Kirkpatrick, 2005, 61-65). But the intensity in how psychologically close God seems to subjects can also be increased with various deliberate methods. Some religious communities in fact teach their members to employ specific methods to obtain a feeling of closeness to God and to cultivate and sustain a lasting personal, emotionally rich relationship to God. Tanya

Luhrmann, who has done anthropological field studies in evangelical Christian communities, describes a range of specific ways in which subjects are taught how to connect emotionally with God. As Luhrmann puts it, evangelical Christians learn ways of "practicing the experience of feeling loved by God" (Luhrmann, 2012, 111). The result of engaging in such practices is that God seems more real in the sense that God's loving presence is felt deeply in one's day to day life. Moreover, to have this connection strongly relates to having belief in God (Luhrmann, 2012, 101-131). This illustrates how the seeming reality of God can become more intense as the result of deliberate activities that enhance the emotional connection to God.

We can represent these and other changes in propensity and seeming reality on the control-board. Let us look closer at this strange instrument. Both knobs have 6 levels. Level 1 is the lowest and 6 the highest. We can think of the formation of the TD as their movement towards 6. We can imagine levels on the control-board that indicate a grey-zone and levels that clearly indicate that the subject is a theist or a non-theist. As we turn the knobs towards 6, it increasingly becomes plausible to see the subject as a theist. In other words, the more relevant and therefore real God becomes for S, the more correct it becomes to think of S as a theist.

A "1" for relevance means that if a cognitive process occurs in S, then the likelihood that this leads to processing of a theistic concept is close to 0. A "6" means that S constantly thinks about God. The concept entangles itself in all processes. For example, it latches on to intuitions about agency-detection, social exchange, proximity to caregivers, and the baseness of out-groups *all the time.* This makes God extremely relevant. A "1" for how real God seems means that God does not seem real at all. A "6" means that God seems fully real.

Now let us consider some combinations of these levels. Consider S, who has "6" for propensity and "1" for seeming reality. S engages in thinking that constantly leads to activation of the theistic concept in belief-conducive ways. The concept successfully latches on to processes, and God is all over the place in S's mind. Yet, God is not real to S at all. In light of Boyer (2002) and Barrett and Lanmnan's (2008) thesis that activation of a religious concept in practical situations strengthens an underlying assumption that a given supernatural agent is real, we can expect God to begin to seem real in S. The extreme relevance adds intensity to how real God seems, but we forcibly keep the knob for this seeming reality at 1. This does not seem a likely psychological reality.

We obtain a psychologically more realistic state by adjusting the propensity slightly downwards and the seeming reality slightly upwards. This allows us to envisage a subject for whom God does not seem quite real yet for whom God is relevant. This subject thinks quite often about God but, for example, the psychological proximity to God is too weak for God to

seem fully real. But given the belief-conducive effects of relevance, God will probably begin to seem more real again. This subject is in a grey zone, not clearly a theist or a non-theist.

Consider next S, whose level for propensity is "1" and seeming reality is "6". God seems fully real to S. In fact, this is beyond doubt. God seems as real as anything can seem. But the mental representation of God is completely dormant in S's mind. No cognitive process ever makes God relevant in any specific situation. It is as if we have turned off the systems whose functioning CSR-theories place at the centre of their explanations. S in this example is atypical and does not plausibly have the theistic disposition. But that does not mean this subject is not a theist. If God seems real to S, then S is a theist even though no intuitive thought undergirds this belief. S would simply be a highly atypical theist. The TD is, as I have argued, a sufficient condition for being a theist, and a representative depiction. However, it is not necessary. One could also potentially form belief in God in a more reasoned way than through the relevance-enhancing processes described above. Moreover, frequent intuitive thought about something or an emotional connection to that something is not necessary for it to seem real. I might for example rarely think about Saturn, but this planet nevertheless seems real to me in the sense that I believe in its existence. It is also possible to persist believing that God exists despite never feeling any emotional connection to God. Consider also individuals on the autism spectrum. These individuals have deficiencies in their ability to attribute mental states to other agents, and they tend not to be theists (Caldwell-Harris et al. 2011, Norenzayan and Gervais, 2013, Norenzayan, 2013, 178-180). But we should not rule out that some of them are theists in their own way. This way of being a theist does not involve much intuitive thinking, but it could involve a more logical and reasoned process. We obtain a more representative subject if we raise the propensity-level somewhat. That allows us to envisage a highly reflective subject who rarely thinks intuitively about God in particular situations but for whom God seems real, perhaps akin to how Saturn seems real even though one does not often think about it.

We have seen some extreme cases and some psychologically possible albeit unusual cases. We obtain more psychologically plausible cases by bringing the knob-levels closer together and by adjusting the propensity-knob to a high level. But is there something like a prototypical TD? Given our discussion of how attachment-related thought influences religious thought, we can expect tokens of the TD to be varied in terms of how responsive, loving, or caring God is construed as being. But a general feature we can expect to recur in TD's is that a large proportion of the thoughts about God that arise are about how God relates to the events in the subject's life. Moreover, given the role of intuitive thought, we can expect a prototypical TD to involve high rather than low psychological relevance. Then, we can add the following to

the definition of the TD given above: The TD is a relatively stable feature of the subject such that (a) there is a *high* propensity for a theistic concept to figure as part of cognitive processes in ways that induce theistic beliefs. As the result of this, God seems real.

5. Intuitive, explanatory, and folk-theological beliefs

We now turn to how theistic beliefs arise in subjects who have the propensity we have discussed. In other words, we are not asking how people become theists but how theistic beliefs arise. This question can be related to the second proximate question Tinbergen (1963) describes, which is about how various factors trigger a given behavior on a given occasion. Three types of theistic beliefs are likely outputs of the theistic disposition.

We start with *intuitive theistic beliefs*. Our account of this is brief because we have already dealt with this above when discussing relevance-enhancing process. Recall the example of how S formed the belief that God disapproves of Z. This is an example of an intuitive theistic belief. We saw that such relevance-enhancing processes gradually contribute to the formation of theistic dispositions by making God lastingly relevant. They also achieve something else. They keep occurring in subjects with these dispositions, thereby sustaining such habits and keeping them going over time. We can call a relevance-enhancing episode occurring in a subject for whom God is (already) relevant a "relevance-preserving" episode. The case above involving cheater-detection is just one example. We can also imagine countless other cases whereby the concept of God for example latches on to intuitions related to attachment, inferring the purpose and function of artefacts, or handling interactions with out-group members, thus preserving the relevance of God. This relates to what Boyer (2002) calls the "aggregate relevance" of a concept, its ability to latch onto the functioning of many rather than few systems. The central feature intuitive theistic beliefs have is that in terms of content, they relate strongly to intuitions produced by evolved cognitive systems.[63]

A second type of theistic belief concerns attempts to explain events. Here the notion of God is not an extra layer on top of intuitive thinking but an integral part of how the subject spontaneously explains things. *Explanatory theistic beliefs* are most likely to occur in subjects

[63] That is why I label them intuitive. But since an MCI concept, "God", is involved, some violations of intuitive ontology are also involved. Intuitive theistic beliefs are therefore not completely intuitive (as we saw in our discussion of MCI theory in the previous chapter), but largely intuitive.

who already tend to form intuitive theistic beliefs and for whom God therefore is relevant. Justin Barrett (2004) presents 3 examples, which in my view plausibly illustrate the formation of these beliefs. The two first involve HADD and the third involves intuitive teleology.

In the first example, Barrett asks us to consider S who observes a storm destroy a house in S's village. This observation triggers agency-detection as the result of how the movements of the storm indicate purposefulness and agency. S tacitly infers during the event that the storm has agency and subsequently employs a religious concept to explain what occurred. This leads to a religious belief involving that concept (Barrett, 2004, 33-34, see also Barrett and Lanman, 2008, 116 for a similar example). Let us assume S uses a theistic concept. Then this is a good example of how a profound theistic belief arises from a prior intuitive belief about agency. The belief is something like, "God destroyed x's house." This belief helps S explain an event that seems to require an explanation. The belief S forms involves an *attribution.* Hood et al. (2009, 13-14) describe attributions as attempts to explain events, people's behaviour, or one's feelings and experiences. The use of simple belief-desire psychology for example involves attributions. Those attributions we are interested in presently are more profound because they involve the notion of an invisible supernatural person who is thought to be involved in a natural phenomenon (a storm). That makes them examples of what Hood et al. call *religious attributions*. An ambiguous event that challenges the subject's feeling of "mastery and control" of the situation can trigger such attributions. Hood et al. suggest that this typically happens when *naturalistic attributions*, i.e. attempts to explain things in terms of natural causes and agents, do not seem adequate and ambiguity therefore persists (Hood et al., 2009, 44-45).

The encounter with a seemingly purposeful storm leads to an explanation in terms of supernatural agency. But this does not happen in any subject. We must, in my view, suppose that God already is relevant. In a subject with the theistic disposition, the threshold for explaining events in this fashion is low. Moreover, such explanations often lead to a barrage of further beliefs. In a subject prone to form thoughts of God, this is usually what happens. Suppose the subject forms the belief that God destroyed x's house. We can now expect a barrage of beliefs about the purposes God had in doing this. Did for example the inhabitants of the house somehow deserve it? There is evidence for a tendency to spontaneously explain the suffering of others as something they deserve. This tendency to "blame the victim" relates to a bias in favour of thinking of the world as inherently moral and just. Nicolas Baumard and Pascal Boyer argue that we have an evolved "fairness-bias", which creates the intuitive impression that the world is just. They suggest that this impression induces those who observe seemingly pointless suffering to rationalize and explain it as deserved punishment (Baumard and Boyer,

2013, 272-273, see also Elster, 2007, 269-270). I surmise that when a theistic concept connects with this kind of spontaneous explanatory tendency, the results are beliefs about divine punishment. The initial explanatory belief, "God destroyed x's house", leads the subject to form another explanatory belief, "God destroyed x's house as a punishment for x's sins!"

The next example J. Barrett provides also involves a dramatic event. An explosion occurred in a silo where Doug works. This has trapped him inside the building. Thinking he is about to die, Doug asks God to take him home. But a voice seems to say, "not yet". Doug then feels "invisible hands" lift him out of a second storey window to safety. Doug's experiences in the silo trigger HADD and intuitive thoughts about the presence of an unspecified agent occur. Later, when Doug ponders what happened, he employs a religious concept suggested to him by his friends. This leads to a belief about a specific supernatural agent (Barrett, 2004, 34-35). Here the result is formation of a religious belief where agency-detection initially contributes together with further subsequent explanatory activity in a social context. This is a plausible suggestion, which allows HADD to have an auxiliary role in the formation of some religious beliefs in some contexts (this moderate role for HADD is further emphasized by Barrett and Lanman, 2008 and discussed in Van Leeuwen and Van Elk, 2018). Suppose the concept Doug's friends suggested to him was theistic. Suppose they told him about God and suggested that divine intervention had occurred. Then, we have an example of the formation of an explanatory theistic belief in Doug, "God saved me from the silo".

Notice that in this example, J. Barrett does not seem to describe a subject already prone to form theistic beliefs. Why would Doug need to talk to his friends to obtain a religious concept? Did he not already have one? Let us amend the example and suppose that Doug is prone to think of God. Then we need not talk about friends that expose him to specific religious ideas in order plausibly to describe him as forming a theistic belief. Given that Doug already is prone to form beliefs about God, it is no surprise that he attributes a seemingly miraculous escape to God. Notice also that if we suppose that Doug is prone to form such beliefs already, then we do not need to introduce astonishing experiences of voices talking or "invisible hands" lifting the subject. We can propose a less dramatic example. It suffices that Doug was frightened, had a vague feeling of a presence, or that he managed to resolve a somewhat uncomfortable situation in the parking-lot. Given a sufficiently strong propensity to think of God, Doug is even in the latter case likely to form an explanatory theistic belief. This could be a belief about how God helped Doug by providing psychological strength to deal with the other driver or perhaps by nudging that driver to abstain from blocking him in. Doug then explains how he managed to resolve the situation in terms of divine assistance from God.

The final example concerns teleological inferences. Justin Barrett points out that when we identify bird's nests, footprints, or tools, we intuitively infer agency as part of the spontaneous explanations we form. We do not believe the structure is an agent, but we see agency as somehow being responsible. Barrett argues that encounters with seemingly non-random structures trigger the belief that they are *designed*, leading to a further belief about a responsible *designer*. It is possible, he suggests, that such encounters in some cases lead to the belief that a designer or "creator" of some kind is responsible. This can help reinforce a given religious concept the subject already entertains, facilitating the formation of religious belief (Barrett, 2004, 36-38, see also Kelemen, 2004, for a discussion of how teleological thought can facilitate the formation of religious beliefs). Based on this, it seems possible that a subject with the theistic disposition who encounters something apparently designed appeals to God in an explanation of that design. An encounter with a complex or puzzling structure thus can have the effect that an explanatory theistic belief is triggered in the subject.[64]

Notice that intuitive theistic beliefs and explanatory theistic beliefs are not about the content of *religious teachings*. They are not about how Muhammed flew to Jerusalem on a horse, that the Virgin Mary ascended into heaven, that God is one yet consists of three persons, that God divided the Red Sea, that God is omniscient, that a meeting occurred between the angel Gabriel and Muhammed in a cave near Mecca, that God entered a pact with the people of Israel, or that Jesus revealed himself to his disciples after his death. Do not theists also hold belief in such teachings? It seems that they do. Of course, they do not interpret all the teachings and narratives found in their respective tradition as factual claims, but we can suppose that at least some of them are interpreted that way. We should then ask whether formation of belief in such teachings can become part of the habit I call the theistic disposition.

Scholars in CSR are quite skeptical towards the notion that theology is relevant to belief-formation. Pascal Boyer argues that theologians will never manage fully to instill religious teachings into believers, and he suggests that this is because human minds distort, misremember, and mistakenly add elements to the complicated ideas they involve, and it is because frequent repetition of "obscure material" becomes "tedious" (Boyer, 2002, 326-327, see also Whitehouse 2004, 66 for a discussion of the "tedium effect"). Frank Keil and Justin Barrett suggest that the theologically correct concept of God that subjects use when reasoning

[64] This is not the result of a carefully reasoned argument from appearance of design to God as designer. For example, this explanatory belief is not like the conclusion to the argument William Paley proposed (for a discussion of this argument, see De Cruz and De Smedt, 2015, 62-68). Rather, it is a non-reasoned explanatory belief where usage of the notion of God helps the subject spontaneously explain a puzzling phenomenon.

reflectively is a "hollow" representation that "masks" an underlying much more anthropomorphic concept, which is involved in actual theistic belief (1996, 241). Finally, Robert McCauley argues that no matter how much theologians try to avoid it, non-conscious cognitive processes tend to distort our theology (McCauley, 2013, 207-219).

A central notion related to these claims is that of "theological correctness". The idea is that subjects entertain two types of religious concepts at the same time, which often contradict each other. The first is a reflective concept containing theologically complex ideas and the second is an implicit or tacit concept that involves distorted and simplified versions of such ideas. Which type of concept the mind activates depends on the cognitive task. People activate the former when asked about what they believe. In this case, subjects try to give the theologically correct answer in accordance with the dictates of their tradition. By contrast, subjects activate the second concept when they think intuitively and non-reflectively about these things (Barrett, 1999, Boyer, 2002, 101-103). For example, the concept of God in classical theism depicts God as being omniscient, omnipresent, and without form or location in time or space. Subjects in a context where construing God this way is considered correct will, if asked, try to conform to this conception. But because of the cognitively demanding nature of this concept, the same subjects will, in situations where there is no time to carefully think about God, use a simpler, anthropomorphic concept according to which God has a spatial location and cognitive limitations (Nicholson, 2016, 7-8, Barrett, 1999, Barrett and Keil, 1996).

The thesis the scholars above are getting at seems to be that subjects do not really believe in intellectually demanding religious teachings based on theology. Subjects who live in literate religious cultures where they are frequently exposed to such teachings are of course able to encode them in memory and use them in reflective thought-processes, but their actual beliefs, which are intuitive and situational, are not really about such teachings. In light of the important emphasis on the practical nature of religion (discussed earlier), it seems correct that the majority of theistic beliefs will be about God in relation to something that occurs here and now. But does that mean theists do not form beliefs in religious teachings? We can challenge this contention by looking critically at an underlying assumption it involves, namely that theological teachings in general are quite complicated and hard to grasp.

Let us start by appreciating that the enterprise of theology, the intellectual, hermeneutical discipline devoted to interpreting and defending religious statements and doctrines, in many cases is cognitively demanding. It is a reflective intellectual activity it takes time to master and which requires careful thought. But that does not make all the religious teachings or doctrines theologians interpret and defend equally demanding for subjects to grasp.

Hard intellectual work can result in teachings that themselves are not hard to grasp. So, even if the activity of theology is complicated, not all the output of theology is equally complicated. Theologians (or preachers perhaps) are therefore not doomed to failure when it comes to producing doxastic effects. Rather, they might be successful with respect to some teachings but less successful with respect to other teachings, depending on how complicated or counterintuitive they are. Let us first consider concepts and doctrines that are very hard to instill into human minds, before observing some examples involving much less difficult material.

Hugh Nicholson points out that the concept of God in classical theism is cognitively demanding since it not only violates naïve physics by depicting God as non-physical and without spatio-temporal location, but since it also violates naïve psychology by depicting God as having limitless knowledge and awareness (Nicholson, 2016, 7-8). Nicholson argues further that the central Christian doctrine of the *Trinity* is an even more cognitively demanding idea.[65] According to this doctrine, God has one divine and eternal substance but consists of three distinct divine persons, Father, Son, and Holy Spirit. This for example means that Jesus, the Son, is of the same substance as God the Father, even though Jesus was a historical person in flesh and blood (Nicholson, 2016, 20). Nicholson has analyzed this doctrine with respect to cognitive costliness and inferential potential. He argues that it not only puts strong demands on human cognition but borders on being "cognitively impossible" (Nicholson, 2016, 22). The reason is that this doctrine violates both naïve physics and psychology (the way the classical concept does) *and* the intuitive thought that distinct persons are distinct substances (Nicholson, 2016, 8). The latter, strongly counterintuitive violation is the result of how the doctrine combines the thesis that there is only one divine substance with the thesis that there are three distinct divine persons, thus rejecting the alternative, less counterintuitive views that (a) there are three divine persons who are distinct substances, or (b) that there is one divine substance that expresses itself in three different ways (Nicholson, 2016, 22). Given the cognitively demanding (or even impossible) nature of this doctrine, the above assumption that theologically correct concepts or teachings are cognitively complex and demanding therefore seems correct in this case (see also Pyysiäinen, 2009, 123-133 for further examples of cognitively demanding religious teachings). This indicates that beliefs that get the *Trinity* right are unlikely outputs of the theistic disposition.

[65] According to Nicholson (2016), the historical development of this doctrine is within the scope of psychological explanation. To show this, Nicholson provides an account of how this doctrine developed (historically) from less demanding views that depicted Jesus as subordinate to God the Father. In this account, an in-group/out-group dynamic plays a central role.

But not all theological teachings are as complicated and difficult as this. Consider the Christian teaching that Jesus rose from the dead. How hard is this teaching to instill in human minds such that they get it right when *not* reflecting deliberately? This teaching does not massively violate intuitive ontologies. The resurrection involves a *breach* with intuitive biology since it depicts Jesus as having bodily functions after death. The resurrected Jesus is for example not a bodiless spirit but a bodily, corporeal person who the disciples can even touch. However, Jesus is still a person with many properties consistent with intuitive psychology. In light of this, the concept of the risen Jesus is not massively counterintuitive but seems instead to be close to Boyer's (2002) cognitive optimum. Yet, it is a theological teaching in Christianity. Developing and refining this doctrine and combining it with other Christian ideas was a demanding task (see Nicholson, 2016). The resulting teaching about a risen Jesus, however, appears easy to grasp.

Consider also the teaching about Muhammed's encounter with the angel Gabriel in a cave near Mecca. According to this teaching, as depicted by Karen Armstrong (1999, 160-163), this angel appeared to Muhammad, challenging him to recite the words of Allah, words that would later become the Quran. The notion of an angel who talks to a human being is counterintuitive, but it does not seem to be massively counterintuitive. For example, there does not appear to be violations of naïve psychology. Yet, this is a central theological teaching in Islam. This illustrates how not all theology is extremely complex and hard to grasp. Consider also the teaching that Muhammed flew to Jerusalem on a horse. A horse that can fly is surprising. It goes against what we learn from experience. But it is not clear that it violates innate intuitions. For example, this notion does not involve a breach with naïve biology, which does not consist of intuitions about which species of animal can or cannot fly. Furthermore, no *transfer* from another domain seems involved. By contrast, if the teaching had described a flying horse that could speak, then it would have involved a violation, a transfer from psychology to biology. But notice that this would hardly have made the teaching too demanding for us to grasp when not reflecting deliberately. None of these teachings are on a par with the doctrine of the Trinity, yet they are also examples of theology. In light of this, not all theology and not all religious teachings are highly counterintuitive or complex. It then seems possible that subjects already prone to think of God, living in a theistic social and cultural setting where there is exposure to these teachings and stories, might eventually form non-deliberate and non-reflective belief about them. In contrast, when it comes to teachings that are quite complicated, such as the one about the Trinity, it seems correct that grasping them and getting them right, if

that is even possible in all cases, requires hard intellectual work. Habitual formation of belief in these teachings seems quite unlikely.

But is the upshot of this that subjects with the theistic disposition simply cannot form *any* belief about the Trinity and equally difficult theological doctrines? I believe we can challenge this. Here is another possibility, which allows us to see how beliefs in both difficult and simple theological teachings can arise: Subjects form belief in a distorted and simplified version of the former type of teaching and a correct belief in the latter type of teaching. Let us explore this possibility.

Even though beliefs about complicated theological teachings are distorted or simplified, their content nevertheless reflects exposure to theology in ways our two prior examples of theistic beliefs do not. In other words, something *is* instilled into people's mind when they form incorrect, distorted, or simplified beliefs about theological teachings. Furthermore, given that not all religious teachings are very hard to grasp, it seems possible for belief in them too to arise. Support for this comes from Barrett and Lanman, who argue that evolved cognitive biases strongly constrain what sort of religious beliefs people are likely to form, but that it is also possible for subjects who receive the right cultural input and who participate in the right religious practices to "stretch" these constraints in ways that allow them to form theologically salient beliefs (Barrett and Lanman, 2008, 118). However, they also point out, strongly counterintuitive beliefs are unlikely to form since that would stretch these constraints too much. Moreover, the beliefs of subjects who fail to engage sufficiently in the culturally specific, "cognition-stretching" practices will not strongly reflect the tenets of theology (Barrett and Lanman, 2008, 119-120). In light of this, we have grounds to introduce a third category of theistic beliefs that arise from the theistic disposition. Let us refer to these as *folk-theological beliefs.* The content of these beliefs strongly reflects exposure to religious teachings. These beliefs either involve a simplifying distortion of complex teachings (that God is a Trinity) or a correct representation of simple teachings (that the angel Gabriel appeared to Muhammad). How can a propensity to form such beliefs become part of the theistic disposition? Here are two possibilities.

Familiarization achieved through repetition of religious materials in a religious social setting can contribute to the formation of folk-theological belief. Through familiarization, theological ideas or teachings that to outsiders seem strange can enter the flow of automatic belief-forming processes in theistic subjects and eventually begin to seem plausible or even obvious. What seems true in the case of highly counterintuitive or complex theological teachings is a distorted and simplified version. For example, subjects might form belief in God

the Father, Son, and Holy Spirit but without getting the exact doctrines of co-substantiality, eternity, and the distinctness of the three divine persons right. By contrast, what seems true in the case of relatively simple and straightforward religious teachings is roughly what these teachings say. For example, by forming folk-theological beliefs, Muslims will get the teaching of Muhammad's encounter with Gabriel right and the same seems possible for Christians with respect to the teaching about the resurrection of Jesus. In both cases, familiarization through repetitive exposure strongly contributes. How do people become familiar with these teachings?

Consider what Harvey Whitehouse (2000, 2004) calls the "doctrinal mode of religiosity". This mode arises in contexts where religion is institutionalized in the sense that there is a hierarchical political structure with a religious leadership, a set of official religious teachings that help mark the identity of the group, and frequent repetition of rituals whereby such teachings are rehearsed. In this mode of religiosity, skilled public speakers who have high social status *repeatedly* expose subjects to carefully formulated doctrines (Whitehouse, 2004, 65-67). J. Barrett chimes in, pointing out how those skilled in political propaganda deliberately use repetition as a tool to create attitude change. He argues that something similar occurs in religion. Merely being told the same religious messages repeatedly for years can lead to acceptance of such messages in the absence of good evidence (Barrett, 2004, 69). A central reason frequent repetition helps instill belief is that it enhances explicit memory (i.e. memory the subject can access). The teachings are encoded in semantic memory and their repetition helps preserve them in the mind (Whitehouse, 2004, 64-66). Whitehouse also describes another possible mechanism that facilitates belief in doctrines. This mechanism is related to how frequent performance of rituals create implicit, procedural memories about how those performances should be conducted. To the extent that repetition of the elements of such rituals over time becomes habitual to the subjects, or as Whitehouse puts it, that subjects are "on autopilot" during these rituals, their tendency to ponder and reflect what the rituals mean or why they must be performed decreases. This creates a perfect atmosphere in which those in power can promulgate the teachings they officially authorize (Whitehouse, 2004, 68).

The doctrinal mode of religiosity is clearly a feature of the Abrahamic religious traditions, which are strongly hierarchical and characterized by a set of official teachings. Moreover, there are frequent, repetitive performances of religious activities of various kinds. These provide many opportunities for exposing subjects repeatedly to the official teachings. Here are some examples. In the Abrahamic traditions, subjects hear, read, or say the same things repeatedly for years in public religious settings where other subjects also do the same. Perhaps the most obvious example is participation in a religious activity such as a ritual in the

synagogue, church, or mosque. Repetition of various elements of a religious tradition is common. Islam requires subjects to perform the same prayers five times a day. In Orthodox Judaism, recitation of the Torah or other texts occurs regularly, and subjects must pray 3 times a day. Furthermore, Judaism features various annual religious celebrations that involve repetition of the same prayers or readings. In Lutheran churches, believers recite the Lord's Prayer every Sunday. They might also frequently make a confession that starts, "Forgive me Lord for I have sinned".

Monastic Christian orders provide striking examples of repetition, which differ from the prior examples in virtue of how the agents themselves *deliberately* attempt to create lasting mental changes in their own minds. By contrast, in the prior examples, the agents do not deliberately try to bring about belief by listening to the same messages repeatedly. Rather, belief-formation occurs as a side-effect of their repetition (in minds already prone to think of God). But in the case of monasticism, the situation is different. Gavin Flood describes practices through which Christian monks or nuns attempt deliberately to change their habituated thought-patterns in lasting ways to achieve the result that their minds in content (what beliefs they have) and in functioning (how they think) fully reflect the tenets of their religious tradition. He describes monasticism as a systematic attempt to re-structure the whole self so that the person in a remarkable sense embodies or manifests the tradition. Such attempts include having to perform hundreds of prayers a day for years (Flood, 2004, 178-181). Tanya Luhrmann's (2012) account of how evangelical Christian communities in the United States teach their congregants to "hear the voice of God" provides another example of how the use of particular practices involving a lot of repetition gradually can create lasting mental changes. In particular, subjects who are new to these communities are taught to reinterpret some of their own thoughts and ideas as being subtle messages from God. Regular, frequent prayer as well as regular Bible-reading sessions are important parts of this practice (Luhrmann, 2012, especially pp. 39-41).

These examples illustrate how the religious life in the Abrahamic traditions typically involves a lot of repetition. Given Whitehouse and Barrett's claim about the role of repetition in belief-formation, these examples provide grounds to conclude that subjects in these traditions will form folk-theological beliefs. The practices they perform, especially the deliberate ones, seem able to "stretch" cognition some distance into theology. This means that in addition to forming intuitive, situational beliefs about God, subjects will form folk-theological beliefs: beliefs that distort teachings that are complicated and counterintuitive or that get simple and less counterintuitive teachings roughly right.

Familiarization through repetition is not the only possible mechanism by which folk-theological belief-formation can become habitual. Here is our second example. It seems that intuitive thought about one's own behavior in a religious setting also contributes to instilling belief. Justin Barrett suggests that merely carrying out activities in religious settings can influence subjects to believe theologically significant things related to these activities. This is because in general, acting *as if* something is true in a public setting can influence subjects to believe that *it is true* (Barrett, 2004, 61-62). There is evidence that theists sometimes act as if things they are not yet sure of is true. As Tanya Luhrmann's study shows, in evangelical Christian communities, subjects are sometimes admonished to deliberately act as if God is present in a very direct and straightforward sense, even if they do not presently experience God in that way, and they might even receive instruction manuals that indicate what methods should be applied (2012, 72-74). A particularly striking example is that subjects set an extra plate at the table for God (*sic*.) (Luhrmann, 2012, 74-75).

Why would acting as if something is the case induce belief that it is the case? We can relate Barrett's statement to a thesis in social psychology, described by Myers (2002), according to which *attitudes follow behavior*. To be more specific, the thesis is that acts carried out voluntarily in public whereby we give assent to statements, or act as if we believe in these statements, make us feel committed. This contributes to making us believe in these statements. A mechanism related to this thesis is dissonance-reduction. A felt discrepancy between what we do in a seemingly voluntary fashion and therefore feel responsible for, and what we believe (related to these acts), leads to psychological tension. By adjusting beliefs (i.e. attitudes) to match the actions one has carried out, the tension is reduced. For this reason, attitude follows behavior. In one study done to test this thesis, experimenters found that subjects who were payed a very low amount to pretend that they advocated for a given (randomly selected) policy subsequently changed their attitudes in favor of that policy. Subjects who were payed a high amount to do the same did not. Scholars have suggested the following as an explanation. The former but not latter subjects were payed a low amount and therefore felt a sense of responsibility for their advocacy. This made them feel committed. In contrast, the latter subjects could explain their behavior in terms of the external reward (the high payment). Therefore, there was no internal tension and no need to change attitude. The former lacked such an explanation and therefore, to offset the tension created by the discrepancy between their behavior and attitudes, adjusted their attitude towards accepting the policy (Myers, 2002, 147-152).

I surmise that something similar happens in religious public settings where subjects in a seemingly voluntary fashion do or say theologically non-trivial things that they in fact have doubts about or at least do not feel fully convinced of. Consider for example things they say during ceremonies or rituals ("Forgive me for I have sinned") or in less structured situations such as in a Bible reading group or at a revivalist Christian rock concert. As long as the subjects do not feel that anyone forced them to say these things, they will feel responsible and committed to what they say. Then, if the above thesis is correct, by dissonance-reduction their initial doubts about the things they say weaken. In this fashion, they gradually accommodate into their belief-structure an increasing amount of religious teachings they previously did not believe in. This seems possible with respect to distorted versions of complicated teachings or correct versions of simple teachings, and thus further supports the notion that theists will, if carrying out the right belief-conducive activities, form folk-theological beliefs in addition to intuitive and explanatory theistic beliefs.

We have seen how a propensity to form folk-theological belief can be integrated into the theistic disposition, which also produces intuitive theistic beliefs and explanatory theistic beliefs. In light of what the current section has argued we can conclude that given the truth of CSR-scholarship, on which our discussion has been based, CSR currently explains how, in general, theistic beliefs arise. Given that we can see individual episodes as tokens of general processes, such an explanation contributes to accounting for individual cases too. It does this by providing general knowledge of belief-formation which together with additional information about the cases in question provides an explanation of how theistic beliefs arise in those cases. Together with the conclusions in section 3 concerning how people become theists, this strongly undermines the weak readings of P3 in E-ENT.

Conclusion

We now conclude the chapter and part 1 of the investigation. In the present and the previous chapter, we have seen how CSR-scholarship explains how people become theists and how theistic beliefs arise. This supports the explanatory adequacy thesis. A general susceptibility to cognitively attractive concepts can in theistic cultural settings gradually turn into a profound belief-forming habit that revolves around the notion of God, a culturally distinct deity related to the three Abrahamic religions. To become a theist is to gradually become prone to think intuitively of God, and to use God as part of a spontaneous explanatory activity. Eventually, if belief-conducive practices are performed, this habit also involves the formation of folk-theological beliefs based on teachings subjects are exposed to in their particular context. These

beliefs either consist of distorted versions of complicated teachings or correct versions of simple teachings. We first attacked the notion that CSR in principle cannot explain how people become theists and how theistic beliefs arise. We then attacked the claim that this research does not presently explain this. Moreover, we argued against the contention that CSR does not contribute to explain individual belief-forming episodes. In light of our conclusions, we can now reject P3 in E-ENT. This central, externalist argument fails to show that CSR is epistemically neutral.

In the next part of the investigation we turn our focus to what normative implications CSR-scholarship has for what I call theist religion. We begin this investigation in the next two chapters, which look at normative implications for theist religion in sense (2). If CSR-scholarship is relevant in assessment of philosophical arguments or theses connected to theism, then it has such implications. We here encounter the internalist argument for the epistemic neutrality thesis (I-ENT). This argument rejects the claim that facts about how beliefs arise are epistemically relevant. Moreover, it contends that CSR fails to shed light on factors internalism recognizes as epistemically relevant. I shall argue against these premises.

Chapter 4. Challenges for Internalist Epistemology of Religion

1. Introduction and chapter-outline

The present and the next chapter look at what relevance CSR-scholarship has in assessment of philosophical arguments or theses connected to theism. The present chapter deals with *theistic evidentialism* (TE) and the next deals with *reformed epistemology* (RE). These both involve a theory about what kind of factors that determine the epistemic status of beliefs, and they make claims about the status of theistic beliefs in light of that theory. As we will see in these two chapters, input from CSR-scholarship is relevant in assessing these positions but bears the strongest on reformed epistemology.

Here is an outline of the present chapter. Section 2 introduces TE. We shall see that epistemic internalism figures as the first part of TE and occurs as premise in the internalist argument for epistemic neutrality (I-ENT).[66] Section 3 argues on the basis of general philosophical considerations (unrelated to CSR) that we should reject epistemic internalism and instead adopt an externalist conception of epistemic justification. This is important because in light of such a conception, our findings in the two previous chapters have potential epistemic relevance. Furthermore, we now have reason to reject TE and I-ENT. In the second part of the chapter, I consider another central part of TE, the claim that there is good evidence for the existence of God. I do not attempt an appraisal of purported evidence for God in general. I instead single out the topic of *morality*, which CSR in my view tells us something important about. Section 4 discusses an argument John Teehan presents, according to which a CSR-based understanding of evolved morality creates what he calls a "cognitive problem of evil". This argument reveals challenges for attempts to incorporate findings of CSR into a theistic framework. Section 5 looks at moral arguments for the existence of God in the light of CSR and discusses a version that Richard Swinburne defends. This version is problematic because it involves a false dichotomy where we seem forced to choose between a theistic account of morality and an account largely based on neo-Darwinian theory. This dichotomy is false

[66]This is an argument for the epistemic neutrality thesis (ENT), the thesis that CSR has no epistemic relevance for theistic beliefs. See chapter 1, section 5 for details.

because there are also many non-theistic philosophical attempts to account for morality that deserve attention.

2. Theistic Evidentialism and the Internalist Argument for the Epistemic Neutrality Thesis

Epistemology of religion is an area in the philosophy of religion that deals with the epistemic status of theistic and other religious beliefs. A central topic in this area is the seriousness of various objections to the justification of such beliefs. According to an evidential objection, there is insufficient evidence for the existence of God to justify belief in God. Alvin Plantinga, who we look closer at in the next chapter, refers to such an objection as "enlightenment evidentialism". A central idea in this objection is that it is only justified or rational for us to believe propositions supported by sufficient evidence (Plantinga, 1983, 24-25). Given such a requirement, the justification of belief in God hinges on whether there is sufficiently good evidence for the proposition that God exists. Plantinga rejects the view of justification involved in this objection (in the next chapter, we will see on what grounds). In contrast, TE, which the present chapter is about, provides a different response to the evidential objection. Advocates of TE agree with (evidential) critics of religion that evidence is necessary for justification. They respond to the objection by maintaining that good arguments provide sufficient evidence for the existence of God to justify theistic belief. Michael Bergmann (2010, 697) defines theistic evidentialism (TE) as the thesis that theistic belief is justified and that it can be so only on the basis of good arguments. TE thus is the conjunction of the following two claims:

> Claim 1: Only evidence provided by good arguments can justify theistic belief.
> Claim 2: Theistic belief is justified (by good arguments).

Claim 1 is a highly restrictive evidentialist thesis about epistemic justification, which only recognizes evidence from good arguments as something that can justify belief in God. We will below look at a more permissive and therefore plausible evidentialist thesis, which allows also non-propositional evidence (that is, evidence not based on arguments). Based on that thesis, we can propose an amended and slightly more permissive version of claim 1, according to which only good evidence can justify belief in God. I argue that also this amended version of TE is

problematic. According to claim 2, good arguments provide sufficient evidence for the existence of God to justify belief in God. In an amended version, the claim says that theistic belief is justified by the evidence. We can plausibly interpret this as the claim that *on balance*, the evidence justifies theistic belief. Advocates of TE have many arguments to choose from as sources of evidence. Consider for example the cosmological argument, the ontological argument, the argument from design, or the moral argument, the latter of which we consider below. They must also deal with purported counter-evidence in the form of arguments that purport to establish that there is no God, such as the argument from evil.

The primary disagreement between theistic evidentialists and evidentialist critics of religion concerns such arguments. Theistic evidentialists believe and the evidential critics deny that arguments establish sufficiently good evidence for the existence of God to justify theistic belief. A typical exchange between them takes the following form: the advocate of theism proposes arguments to show that God exists and then the critic of theism attempts to show that these arguments fail. That critic might in addition - though this seems optional given the assumption that the theist has the burden of proof - propose arguments against theism to bolster his or her case. Richard Dawkins, a biologist who promotes atheism, follows this strategy. Dawkins (2007) criticizes arguments for theism and provides an argument for atheism. Interestingly, Dawkins also includes CSR-scholarship into his attack on religion. The result is a thought-provoking hybrid evidentialist and CSR-based attack. In chapter 8, section 3, we shall see why this, at least in principle, is a legitimate strategy to pursue for an advocate for atheism.

Let us also consider an exchange between TE-advocates and those, like me, who argue that research in CSR *has* epistemic relevance with respect to theistic beliefs. The former argues against the latter that only evidence from arguments and perhaps non-propositional evidence can justify beliefs. Furthermore, CSR, a study of something else than such arguments or such evidence is irrelevant with respect to the epistemic status of belief in God. Instead, this line of argument suggests, only a thorough discussion of the quality of evidence in favour of theism can establish important conclusions concerning such an issue. It is therefore simply a mistake to engage with CSR to settle it. Those who argue this way might view something like the kind of exchange I described above between the TE-advocate and the evidentialist critic as the right way to proceed in a discussion of whether we should believe in God. Who argues in this way? As we saw in chapter 1, section 5, Joshua Thurow, Jonathan Jong, and Aku Visala argue that the central issue is what evidence (or reasons or arguments) theists have in support of their

beliefs, and that CSR does not tell us much (if anything) about this issue.[67] At play in this kind of argument, which attempts to show that such research has no (or very limited) epistemic relevance for theistic beliefs, is a thesis known as *epistemic internalism*. Here is a reminder of the internalist argument for the epistemic neutrality-thesis (ENT) described in chapter 1.

I-ENT

P1: Only factors that are internal to the subject have epistemic relevance.

P2: CSR does not shed light on internal factors (like reasons, evidence, or arguments).

C: CSR has no epistemic relevance for theistic beliefs (ENT).

The next section argues that the internalism in this argument and in TE faces objections, and that in light of these, we should instead endorse an *externalist* view of epistemic justification, thereby denying premise P1 in I-ENT and claim 1 in TE. Then, in the second part of the chapter we see how CSR is relevant for some specific arguments for the existence of God. This strongly undermines P2 in I-ENT, which claims that CSR sheds no light at all on such factors. Our findings will also, but to a lesser extent, undermine claim 2 in TE, by establishing that at least *some* arguments on which theists *might* base their beliefs are problematic.

3. What kinds of factors have epistemic relevance? The internalism/externalism debate

Internalism is the view that all epistemically relevant factors are internal to the subject. Externalism is the denial of this view and therefore the position that *not all* epistemically relevant factors are internal (Goldman and McGrath, 2015, 42).[68] There are two central views on what it means for a factor to be internal in a sense relevant to justification. *Accessibilism* is the view that the internal is what the subject can access cognitively or introspectively.

[67] As I mentioned in chapter 1, section 4, Visala (2011, 160) grants that CSR might undermine the argument from the phenomenon of widespread belief in God to the conclusion that God exists. Visala therefore does not believe that CSR literally has no relevance for any theistic argument. Nevertheless, in the light of statements he and Jonathan Jong make, illustrated in chapter 1, section 5, it seems fair that they endorse P2 in I-ENT or something close to it, like the thesis that CSR does not shed light on any central or important argument for theism.

[68] In chapter 6, I defend a specific version of externalism known as "agent-reliabilism", which holds that it is not necessary for the justification of a belief that internal factors are operative. Instead, it is necessary and sufficient that cognitive abilities are relevantly involved in its production (Greco, 2007, 2012).

According to *mentalism*, the internal means simply the mental (Conee and Feldman, 2011, 54-55).

Notice how internalists can appeal to these views in a defence of the epistemic neutrality of CSR with respect to theistic beliefs. In limiting the range of factors that are epistemically relevant, both versions of internalism seem to support the contention that the epistemic relevance of CSR is limited. CSR-research largely is about cognitive events that *underlie* religious belief and behaviour. It is a central thesis in this research that a large part of the cognition relevant to understanding religion is non-conscious and automatic and thereby outside the domain of what an accessibilist deems epistemically relevant. In relation also to mentalism, it seems that CSR has a limited ability to tell us about epistemically relevant factors. This is somewhat less clear because the domain of the mental seems more difficult to pin down given how Conee and Feldman (2011) describe it. For example, they talk about the mental as those factors that are "internal to the mind" or that "occur within cognitive beings" (Conee and Feldman, 2011, 57). Non-conscious cognitive events inaccessible to introspection are internal to the mind at least if we go along with the view that large parts of the mind are opaque to the subject. The cognitive psychology to which CSR-scholars typically appeal clearly takes this view of the mind. Consider for example Pascal Boyer's discussion of the importance of cognitive processes that we cannot monitor or control (Boyer, 2002, 106-109). This discussion clearly is about what goes in "within cognitive beings". If the mental simply is the domain of what occurs inside of cognitive beings, that is, their brains and information-processing systems, then it seems CSR can tell us something about the mental after all. However, when we look closer at the concrete examples of justifying factors Feldman and Conee envisage, it becomes clear that they have a rather restrictive view of what counts as the mental, which makes their view resemble what accessibilists endorse. We see that the "mental" largely amounts to human experiences and beliefs. Moreover, Conee and Feldman argue in favour of the view that *only* the mental (in this sense) can be evidence. This means that facts about the structure of the universe, the nature of human morality, non-conscious cognition, footprints in the mud, a good argument for the existence of God, and traces of DNA on a piece of fabric are not evidence. Rather, experiences and beliefs that involve awareness of such things can be evidence.

Let us look closer at internalism, starting with Conee and Feldman's mentalism, which provides a representative example of internalism today.[69] Mentalism holds that the normative status of beliefs depends entirely on conscious mental states (herafter: "mental states"). For

[69] For present purposes, I understand internalism to be equivalent to this view and accessibilism and mentalism to amount to the same thing. For a discussion of these and related views, see Prichard (2011, 235-236).

example, the factors that determine the epistemic status of S's belief that p at a specific moment are the mental states S is in during that moment. Conee and Feldman present their view as a supervenience thesis about normative status. According to this view, the epistemic status of a subject's beliefs at time t supervenes strongly on that subject's mental states at t. This means that mental states that occur at the time of belief completely determine the epistemic status of belief. An implication of this is that two mental duplicates, i.e. two mentally identical subjects, must also be epistemically identical. In other words, it is not possible for two mentally identical subjects to be epistemically different (Conee and Feldman, 2011, 54-56).

Let us now consider evidentialism, which is a particular mentalist view of justification and knowledge defended by Conee and Feldman. Evidentialism is the thesis that a doxastic attitude D towards a proposition p at time t is justified for S at t if an only if D is an attitude that fits the evidence S has at t (Conee and Feldman, 2011, 83). As Conee and Feldman point out, evidentialism has been criticized on the grounds that it leads to an intuitively wrong verdict when it comes to subjects who have good evidence for p but fail to use it as a basis for believing that p. For example, one might have good evidence in favour of a belief that p but base the belief that p on irrelevant considerations that do not support that belief. In such a case, it seems wrong that one's belief that p is justified. If so, then evidence seems insufficient for justification. In response to this objection, Feldman and Conee maintain that an attitude of belief towards a proposition p is in fact justified by the evidence one has in a specific sense of the term "justification". In order to understand how they use this term, we need to consider another term Conee and Feldman also introduce, that of "well-foundedness". Well-founded beliefs are attitudes that fit the evidence one has and are based on that evidence. This means such beliefs are not only doxastic attitudes that would be justified in the light of a piece of evidence, in their sense of "justification". Rather, well-founded beliefs are doxastic attitudes subjects have that arise in a proper fashion, namely in a way that involves basing them on justifying evidence. With this distinction, Conee and Feldman respond to the above objection by arguing that a subject who fails to base his belief that p on evidence lacks well-founded belief (Conee and Feldman, 2011, 92-93).

Well-founded belief is according to Conee and Feldman a requirement for knowledge. That is, knowledge that p requires well-founded (true) belief that p. If one has evidence that justifies an attitude of belief towards the true proposition that p and bases a true belief that p on this evidence, then one knows that p (Conee and Feldman, 2011, 104-105). It is clear from this that well-foundedness plays the role in Conee and Feldman's theory that the term "justification" plays in many other theories in epistemology, namely being one of the requirements true belief

must meet in order to be knowledge. Well-foundedness therefore refers to what often is called *doxastic justification*, which means the justification of a belief a subject has. When we say that S's belief that p is justified or that S in a justified manner believes that p then we refer to doxastic justification. In that case, we are evaluating a belief a subject has. This is something other than what Feldman and Conee mean with their term "justification". What they indicate with this term is what often is called *propositional justification*. This involves relations between a subject, evidence, and a proposition. Justification in this sense does not apply to beliefs subjects have. Propositional justification is about the situation that a subject is in with respect to a proposition in light of evidence, for example the situation of being entitled to justifiably believe it, suspend judgement, or disbelieve it. That S is propositionally justified in believing that p does not entail that S believes that p. It means that S is justified in having the belief that p (whether or not he or she actually has this belief) (Dougherty, 2011, 11, Goldman and McGrath, 2015, 9, 26).

Let us digress briefly to relate this distinction to our overarching discussion of theistic beliefs. If for example S has good evidence for the existence of God (suppose S is aware of a very strong argument for theism), then *theistic belief* is propositionally justified for S. This could be true even if S is an atheist (or an agnostic). An atheist is propositionally justified in believing in God if strong evidence this atheist has supports the proposition that God exists. TE (the conjunction of the two claims specified in section 2) says that arguments establish sufficient evidence for the existence of God to justify the belief in God. It is thus a thesis about propositional justification. It is not a psychological thesis about how theists form their beliefs or about their doxastic justification. If TE is true, then (given Conee and Feldman's version of internalism) subjects aware of justifying arguments for theism are propositionally justified in believing in God. Furthermore, well-founded belief in God is something subjects then *can* obtain by forming such a belief *based on* those arguments. Notice that even if all of this is true, it is still far from clear that theistic beliefs in representative cases are well-founded. Recall my discussion of confabulation and about three types of theistic beliefs that arise from theistic dispositions. In light of this discussion, strongly informed by CSR-scholarship, we have strong grounds to deny the view that theists usually base their beliefs in God on arguments. Then we can also deny the view that theists usually meet necessary requirements for well-foundedness in accordance with evidentialism. We can deny this *even if TE is true* and good arguments for theism therefore exist. This indicates that evidentialism is a restrictive theory with potentially

negative implications for many religious beliefs.[70] In contrast, externalism opens up for beliefs to be (doxastically) justified also without being based on evidence (see chapters 5 and 6 for more) and is therefore more permissive. Let us now return to our main task, which is to show that internalism is *not* correct and that we therefore have strong grounds to *deny* that TE is true.

We will now see that mentalist internalism and in particular evidentialism face strong challenges. I will present a variety of cases. The first targets internalism, arguing that epistemic status does not supervene on mental states: there are cases where subjects are mentally identical yet seem to be epistemically different. The second case argues against evidentialism by attempting to show that evidence is not sufficient for (doxastic) justification. The third argues that evidence is not necessary for (doxastic) justification. The fourth and final case points to a challenge for internalists when it comes to understanding the basing-relation. This relation is involved in the notion of well-foundedness, which is central in evidentialism. Understanding the basing-relation is difficult without introducing the ideas of a psychological and causal process. Given its internalist commitments, evidentialism should not open up for such ideas to have much relevance, but it is hard to make sense of well-foundedness without doing so.

Case 1: Variation in extra-mental facts produce variation in epistemic status

If epistemic status supervenes on mental states, then two mentally identical subjects cannot be different with respect to the epistemic status of their beliefs. In other words, mental duplicates must be epistemic duplicates. Feldman and Conee defend this mentalist supervenience claim by presenting pairs of cases where it seems obvious that the subjects are not justificationally identical. They then appeal to a purported mental difference as a best explanation of why that normative difference obtains. For example, S and S* both believe it is warm today. Both have read a weather forecast that says it will be warm but S* has also gone out and felt the heat. The experience S* has makes S* more justified than S in believing it is warm. What accounts for this difference is a mental difference. Another example concerns a subject who is a brain in a vat induced to believe falsely that he had breakfast this morning. In one scenario, stimulation of his brain induces this belief through creating a faint seeming memory-experience. In another scenario, stimulation triggers what seems to be a strong and vivid memory of having breakfast this morning. There is an epistemic difference between the subjects in these two scenarios and a mental difference between them provides the best account of it. Based on these and other similar examples, Conee and Feldman suggest we draw the conclusion that *only* changes in

[70] To be more specific, I suspect - but will not attempt to show - that such implications obtain given internalism and my account of belief-formation in chapter 3.

mental states bring about changes in epistemic status. This means for example that changes in the etiology of belief or the environment do not make an epistemic difference (Conee and Feldman, 2011, 58-61). Conee and Feldman point out that if internal (mental) differences cannot fully account for epistemic differences in a given case, then the theory has a problem (2011, 78).

In response to this, we should not argue that differences in mental states do not bring about differences in epistemic states. Rather, we should critically scrutinize the claim that *only* differences in mental states entail epistemic differences. In particular, if we have reasons to believe that at least some epistemic differences obtain between mentally identical subjects, then we have reasons to believe that internalism is wrong. Do any variations in extra-mental factors plausibly entail epistemic differences? Consider an argument John Greco (2010) presents. He describes pairs of cases where subjects are mentally identical and extra-mental facts vary and then goes on to argue that we have good reasons to view the subjects as epistemically different because of these extra-mental differences. One of his examples involves facts about etiology of belief. Consider two subjects who both know a set of logical axioms and inference-rules. Their knowledge of logic is identical. The first subject uses his knowledge of logic when reasoning correctly to a true belief and the second subject fails to use this knowledge and reasons in a flawed manner to the same true belief. The subjects lack introspective access to the reasoning-procedure their minds used when pondering the problem and the differences in reasoning do not bring with them differences in mental states. Yet, there seems intuitively to be an epistemic difference between them. This means that two mentally identical subjects can be epistemically different (Greco, 2010, 50-53).

We can also consider a case where there are variations in the kind of environment in which two mentally identical subjects find themselves. I base this case on the fake-barn example described by Goldman (1976, 772-773). S is driving down the road and sees a barn ahead of him in broad daylight. This leads to the true belief that he sees a barn. This true belief amounts to knowledge and therefore S knows that there is a barn ahead of him. Another subject S* drives down the road in fake-barn country and sees a barn ahead. The area in which S* drives contains many barn facades that look exactly like real barns from the road. The object S* sees is a real barn but S* could easily have looked at another object and formed the false belief that he saw a barn. Assume that S and S* are mentally identical. Most epistemologists agree that the kind of environment in which S* finds himself in threatens the epistemic status of his beliefs. For example, epistemologists widely agree that the subject in a fake-barn context does not know that he sees a barn even though his belief is true. I therefore take it as relatively uncontroversial

that there is an epistemic difference between S and S*. S knows he sees a barn but S* does not know this.[71] Since these two subjects are mentally identical, mentally identical subjects can be epistemically different. A difference with respect to extra-mental facts brings with it an epistemic difference. In conclusion, these two (and similar) cases plausibly indicate that epistemic status does not supervene on mental facts and this result undermines internalism.

Case 2: Wishful reasoning and forgetfulness do not make beliefs doxastically justified.

Alvin Goldman focuses attention on Feldman and Conee's claim that epistemic status depends only on mental states occurring *at the time of belief.* For example, this means that past mental states are not epistemically relevant factors. Goldman associates this claim with the thesis that an experience at t as of remembering that p confers justification on my belief that p at t. For example, if I now seem to remember that I had an ice cream yesterday, then this justifies me in holding that belief now (Goldman, 2015, 107). Such an experience presumably is the kind of evidence Conee and Feldman envisage as being sufficient for justifying beliefs. Goldman provides the following case in order to demonstrate why this is not plausible: S forms the belief that p at time t* (a time before t) through wishful thinking. At t* this was not a justified belief because beliefs that arise through wishful thinking are not justified. However, at t S has forgotten how his belief arose. S simply has an experience as of seeming to remember that p. This presumably provides sufficient evidence to justify S's present belief that p. Goldman objects that this is not plausible. A belief that was unjustified in the past does not become justified because the subject fails to be aware of how it arose and presently has an experience as of its truth (Goldman, 2015, 108). This negative verdict also applies to well-foundedness. Imagine that the subject in this example bases his present belief that p on his experience of seeming to remember it. This subject then has a doxastic attitude of belief towards p and this attitude fits the (experiential) evidence the subject has. Moreover, S bases the attitude on this evidence. Evidentialism entails that the subject has well-founded (doxastically justified) belief, but this does not seem to be the case. If so, then basing beliefs on evidential mental states is insufficient for well-founded belief.

[71] Notice that both S and S* see a real barn and that relevant facts concerning what they see therefore do not change. This means the difference in epistemic status is not the result of a difference with respect to a fact or the truth-value of the beliefs they form. If for example S saw a barn but S* only saw a barn-façade but the perceptual experiences were identical, then also a mentalist might grant that there is an epistemic difference. What brings about this difference is that one can only know what is true, so if we change what is true, then we also change what people know. This is not the case in the present two scenarios.

Case 3: Forgotten evidence does not remove the justification of a belief originally based on good evidence.

Goldman asks us to consider S who forms a belief that p based on good evidence in the past. Presently, S no longer remembers his evidence for p. S has no subjective experience of seeming to remember the evidence that once justified the belief and no doxastic resources (i.e. beliefs) to appeal to as support. S has retained the belief in memory. Evidentialism entails that S lacks well-founded (i.e. justified) belief. This is because S has no conscious experiences or beliefs that constitute evidence on which he can base his belief. This result is problematic. S's belief that p seems justified. It is too strict to say that justification disappears if awareness of the original justifying factor ceases (Goldman, 2015, 129-130). This shows that factors that count as evidence according to Feldman and Conee, such as experiences and beliefs, are not necessary for well-foundedness.

Case 4: It is difficult to understand the basing-relation from an internalist perspective.

Well-foundedness requires forming a doxastic attitude towards propositions based on evidence that the attitude fits. How can evidentialists account for the basing-relation? Conee and Feldman suggest that basing might involve a causal element but stress that what is central is an "appreciation" of the evidence. Their suggestion is that when S bases his belief that p on evidence e, then this means that S appreciates e when forming the belief that p (Conee and Feldman, 2011, 93). This presumably means that basing is (largely) something mental. For example, there is a mental difference between a subject who bases his belief on a piece of evidence and another subject who is mentally identical to him in all other respects, who does not base his belief on this piece of evidence. However, it is not entirely clear what it means to "appreciate" evidence. Furthermore, we have reasons to deny the view that basing is mental (or largely so). Jack Lyons favours another way to look at basing, where the crucial element of it is that it is a (mechanistic) psychological causal process. When S bases his belief that p on evidence e, then e is part of the psychological process by which S's belief that p arises. This means that the evidence that forms the basis for the belief causes the belief (or is part of the causal process that brings about the belief). Lyons plausibly suggests that such a psychological process is not necessarily conscious and does not necessarily bring with it a corresponding mental state. It is possible, he suggests, to take something as evidence when forming a belief but without consciously being aware of doing so (Lyons, 2009, 138-139). This is a more plausible understanding of the basing-relation but not one to which evidentialists can appeal. This is because evidentialists do not see the non-mental as epistemically relevant, but clearly

recognize basing as relevant. Therefore, it seems that they must appeal to the less plausible view that basing is conscious and largely mental in order to grant it epistemic relevance.

Before concluding let us note that Conee and Feldman acknowledge that in appealing to the notion of well-foundedness, evidentialism to some extent resembles reliabilism, a prime example of an externalist theory. It might even turn out that well-foundedness and some versions of reliabilism are equivalent, they suggest. Moreover, they argue, to the extent that this is true, it should be possible to cash out the intuitions and insights these forms of reliabilism involve in evidentialist terms, namely using the notion of well-foundedness (Conee and Feldman, 2011, 94). From an externalist reliabilist perspective, however, things look different. It seems that Conee and Feldman make a concession to reliabilism and that "well-founded belief" is another term for what a reliabilist views as evidentially justified belief, i.e. belief produced by a reliable process that involves consideration of evidence. As Jack Lyons suggests, it is possible for a reliabilist to see evidentially justified beliefs as part of a larger landscape of beliefs, many of which are non-evidentially justified by processes that do not involve consideration of evidence. This form of reliabilism distinguishes evidential from non-evidential, epistemically relevant factors and allows both (Lyons, 2009, 21-26). Reliabilism readily opens up for psychology to provide insights into the cognition involved when subjects base beliefs on something, and it does not matter whether basing is non-mental or involves mental events.

Even though we should not make final or definitive statements about the contentious and complex internalism/externalism issue in the light of the above cases, they do in my view put pressure on internalism by providing good reasons to accept the thesis that there are epistemically relevant factors outside the domain of the mental. In light of these and similar cases, I am therefore inclined, along with many others[72], to conclude that externalism is correct. As mentioned, this is not the view that no internal factors are relevant but the view that *not all* factors that are epistemically relevant are internal. Externalism does not deny that evidence provided by beliefs and experiences is relevant but denies the strong claim that *only* evidence can justify and that *only* mental states can be evidence. Externalism therefore denies P1 in the internalist argument to neutrality (I-ENT), which only allows internal factors epistemic relevance. Externalism also denies claim 1 in TE, according to which belief in God can only be

[72] How many? The support for externalism is relatively strong in current philosophy. In a survey of what positions professional philosophers accept, David Bourget and David J. Chalmers (2013) found that 42.7 % were externalists, 26.4 % were internalists, and that 30.8 % supported other views.

justified by good arguments, or in the amended version, that this belief can only be justified by evidence. We now conclude that these two claims are wrong. Externalism denies these claims by widening the range of factors that are epistemically relevant to include also other things, such as facts about the environment and non-conscious cognition. Notice that given this widening, failure to base beliefs on arguments (or other forms of evidence) does not entail lack of doxastic justification. This means that it is at least in principle possible for theists to have justified beliefs about God if they lack good arguments or good evidence to support their beliefs. If extra-mental factors related to the environment or the formation of beliefs confer justification, then they are justified also if the subjects lack evidence. *Reformed epistemology*, which we investigate in chapter 5, appeals to externalism to show that theistic beliefs can have positive epistemic status despite an absence of arguments or evidence to support them. The two most important extra-mental factors that according to this theory must be in place are the existence of God and the existence of a god-faculty in our minds (Plantinga, 2000). I argue in chapter 5 that we lack this faculty. Furthermore, I argue against reformed epistemology and in favour of a related but more plausible form of virtue-oriented externalism known as "agent-reliabilism" (Greco, 2007, 2012) in chapter 6. We shall then in chapter 7 see that CSR sheds light on factors that this externalism recognizes as epistemically relevant.

I now turn to the second part of TE, the claim that theistic belief is justified by good arguments (or good evidence). This is a very important issue for advocates for TE because they also maintain that only good arguments (or good evidence) can justify theistic belief. Then it is obviously crucial for the truth of TE that such arguments (or such evidence) obtain. It is also crucial that arguments against theism do not succeed. We can interpret claim 2 in TE as saying that on balance the arguments and evidence favour theism. We shall not investigate theistic arguments (or purported evidence) in general but instead explore what relevance *CSR* has for this issue. As mentioned in chapter 1, section 4, theistic arguments often involve abstract issues that philosophers deal with "from the armchair". CSR does not bear strongly on such arguments in general. Nevertheless, this research is relevant with respect to some arguments, both against and for theism.[73] I have singled out the topic of *morality*. This topic is relevant in a discussion of the problem of evil, attempts to accommodate CSR-findings to a theist philosophy, and moral arguments for theism. Such research therefore has a bearing on theistic arguments and theses. This strongly undermines P2 in I-ENT, which claims that CSR does *not* shed light on internal

[73] An example is the argument for theism from the phenomenon of widespread belief in God (as mentioned in Ch. 1, section 4).

factors like reasons, evidence or arguments, and it undermines claim 2 in TE (to a lesser extent) by creating problems for the thesis that the balance of the evidence favours theism.

4. John Teehan and "the Cognitive Problem of Evil"

It is a common notion in the science and religion debate that science is about how things happen and that religion explains why they happen, and that the two therefore are not in conflict. A theist positive to this way of arguing might suggest that CSR provides insights about the "nuts and bolts" of religion while theistic philosophy provides an interpretation of these findings in terms of divine purposes and goals. Are the insights that CSR provides themselves neutral with respect to the viability of such an interpretative enterprise? Is theism an optional framework those who believe in God can simply add to the findings of science? If yes, then the plausibility of claim 2 in TE, the claim that there is sufficiently good evidence for the existence of God to justify the belief in God, seems unaffected by cognitive research on religion. The theist can argue in favour of theism on cognition-independent grounds and then, at least if the findings of current cognitive science are seen as important and therefore necessary to deal with, integrate those findings into a theistic framework. The present section discusses findings that create a serious challenge for such attempts, and which have the potential to affect the balance of the evidence for theism. We then ask what implications these findings have for theist religion.

In a recent discussion of implications of CSR for religion, John Teehan (2016a) points to how theistic philosophers responded to the emergence of Darwin's theory of evolution by viewing God as designing and guiding the evolutionary process. They thus added a theistic interpretative overlay on Darwin's findings. Presently, Teehan points out, some theistic philosophers and scholars respond similarly to the emergence of CSR-research. As an example of this, Teehan points to Justin Barrett (2004), who suggests that God possibly "orchestrates" evolution to produce beings susceptible to religion. This opens up for thinking that CSR in a sense discovers how God enables belief in him. Teehan refers to this sort of cognitively informed theistic thesis as a "cognitively compatible theism". He argues that the evolutionary and cognitive study of morality (which CSR in a broad sense incorporates) provides a "cognitive science of evil" which seriously challenges such a thesis. The challenge is that such a science leads to a problem of evil. A God who not only creates the world but also actively guides evolution to sculpt our cognition is morally responsible for the evil that arises from the

normal functioning of our evolved cognition (Teehan, 2016a, 41-49). This argument thus points to a serious challenge for attempts to integrate CSR-findings into theistic philosophy, and it is relevant for claim 2 in TE, since the overall balance of the evidence for theism is affected.

What is this science of evil and how does it create problems for theistic theses and arguments? Let us start with the first question. According to Teehan (2010, 2016a, cf. Joyce, 2006, Tomasello, 2016), human morality is an adaptation for handling social existence in groups (see the next section for a closer look at the notion of evolved morality in general). A crucial feature of this adaptation (mentioned in chapter 3) is that it involves a bias towards members of the in-group and against members of the out-group. Evolution seems to have made us empathetic towards in-group members and insensitive to the plight of out-group members. This bias contributes to the cohesion and stability of groups, something that has fitness-benefits for individual members (Teehan, 2016a, 43-49, Teehan, 2010, 41-42, see also Krebs, 2005, 763, Kirkpatrick 2005, Norenzayan 2013, Tomasello, 2016, and Boyer, 2018). We are even liable to dehumanize others, Teehan (2010) argues. Consider words used to describe predators, prey, or parasites. Labelling other human beings with such words can powerfully influence how we think of their moral status. Thoughts about members of an out-group as parasites can for example trigger an evolved system for handling exposure to contagious material. This system produces feelings of disgust upon an encounter with the supposed contagious person. When it is common in one group to label members of another group with terms that trigger this system, the result can be that the former group dehumanizes members of the latter. This can even break down psychological barriers towards engaging in genocide (Teehan, 2010, 156-157). As an example, consider the language involved in the genocide in Rwanda. The Hutu referred to the Tutsi as "cockroaches", obviously a label that indicates the presence of something contagious. According to Teehan, in these and similar susceptibilities related to our evolved morality, we find,

> "the psychological basis for prejudice, discrimination, and dehumanization – and all the injustice, harm, and violence that follow (Teehan, 2016a, 48)."

Importantly, our susceptibility to disregard or dehumanize out-group members is a *design-feature* and not the result of cognitive malfunctioning (Teehan, 2016a). This, in brief, is the cognitive science of evil.

The problem this creates has to do with the moral responsibility God has in sculpting us in such a fashion that so much misery would come about from the normal functioning of our

minds. By providing insights that indicate that God is responsible for so much evil and suffering, CSR leads to a cognitive problem of evil. How serious is this problem? Teehan contends that if God is morally responsible for our evil because our evil is the result of divine design, then "theodicy fails" and we have a "powerful case for debunking" (2016a, 49). CSR then provides powerful evidence against the truth of the theistic thesis that God sculpted our cognition, and in extension seriously undermines theism itself, at least in the Christian version. This is because the viability of theism hinges on whether an adequate theodicy is available, and the cognitive science of evil seriously hinders attempts to produce one (Teehan, 2016a, 41-42).

The problem of evil in general is the well-known challenge of how to reconcile the existence of an omnipotent, omniscient, and perfectly good God with the existence of evil and suffering. It is thus a problem that arises within a theistic metaphysical framework.[74] A theodicy is a response to this problem which attempts to show that God is morally justified in allowing evil to occur (Taliaferro, 2004, 299-300). The common notion of evil in the philosophy of religion is quite broad. It is not primarily used as an adjective to refer to behaviour we find reprehensive. Rather it is a noun which refers to the fact of suffering. First, we have what is known as "natural evils". This means cases of suffering brought about by natural causes, such as disasters and disease. Second, we have "moral evils". This means cases of undeserved suffering brought about intentionally through acts like murder, torture, the creation of unjust societies, and so on (see Taliaferro, 2004, 301, Swinburne, 2004, 236).

The present argument claims that if God sculpts our cognition in such a way that we become susceptible to evil, then God is morally responsible for this. Moreover, he is, since theodicy fails, morally unjustified in doing it. Here is an initial response according to which God *is* morally justified.

God has a morally justifying reason to allow evil and suffering if the presence of these things enables a greater good to obtain. One such good is the existence of humans capable of real moral choices, which requires that we be free. It seems that God cannot hinder us when we choose to do something evil without also taking away our freedom and therefore also the capacity for moral choice. If freedom is a greater good such that God is morally justified in allowing evil and suffering as a (necessary) means to enable it, then theodicy seems to succeed (see Taliaferro, 2004, 305-307, for an account of the greater-good theodicy). Perhaps God is

[74] Taliaferro (2004, 300) refers to it as the "theistic problem of evil". This is the problem we shall focus on. But does evil only conceivably create problems for theists? Yujin Nagasawa (2018) argues that this is not the case: evil creates a problem for atheists who commit to "existential optimism", the view that the world is a largely good place that we should be grateful to exist in. The present investigation does not investigate this argument.

justified in sculpting us the way indicated above because if he instead programmed us to be naturally good, then that would compromise our freedom and therefore also our morality. Does this response effectively handle the cognitive problem of evil? We have grounds to answer in the negative.

We might accept the notion that to enable human freedom, God must abstain from determining our actions. If we cannot perform any evil act because God always intervenes to stop us, we are not free. This lack of freedom compromises our ability to make moral choices. However, it does not seem necessary for God to enable human freedom (and therefore morality) that he sculpts our cognition in such a way that we have a bias towards being *evil*. Rather, it seems sufficient (for our freedom and therefore morality) that we have no bias in any particular direction and that God abstains from intervening when we are about to do something bad. Furthermore, God (who by definition is perfect) could sculpt our cognition in ways that made us susceptible to believe in him without also adding evil to our design. It is not plausible that the only possible way for us to have a cognitive susceptibility to form theistic belief is by that kind of arrangement. Moreover, as I have argued in chapters 2 and 3, we do not have a specific susceptibility to belief in God. Rather, we have a general susceptibility to MCI concepts which in theistic cultural contexts can gradually develop into the theistic belief-forming habit I call the theistic disposition. We now have reason to add the cognitive science of evil to the structure of this susceptibility. The upshot is that our natural cognitive makeup departs strongly from what we can expect it to look like given that a loving God sculpted it. This challenges integration-theses and possibly also theism itself. Let us start with the latter.

As we saw above, Teehan suggests that the cognitive science of evil strongly motivates *debunking*. Is the idea that in virtue of providing a cognitive understanding of evil which is *logically incompatible* with theism, CSR provides conclusive grounds for a rebuttal of theism? This seems close to what Teehan is arguing, given his claim (2016a, 49) that theodicy fails in light of the cognitive science of evil. If theodicy fails, then that provides strong grounds to reject theism. If the truth of claims in CSR entails that theodicy fails, then such findings have strong negative bearing on the truth of theism. We might then even say that CSR and theism are incompatible. But is this necessarily the case? I shall not deal with this interesting question, since I have assumed (Ch. 1, section 3) that CSR and theism are or at least can be logically compatible, and since the main goal in the present investigation is to explore other ways in which CSR has implications for theist religion. Interestingly, implications for theism seem to obtain even on the assumption of logical compatibility. Let us explore how.

The claim that CSR and theism are logically incompatible seems related to the *logical* version of the problem of evil, which involves the claim that theism and the fact of evil are logically incompatible. This means, if correct, that we can deduce the falsity of theism from the fact of evil. By contrast, the *inductive problem of evil* involves the claim that the truth of theism is less probable in light of the massive amount of evil we find in the world. As some for example argue, though a loving God could allow some evils to occur, it is highly unlikely that God would allow the massive amount of evils the world contains (see Taliaferro, 2004, 299 for the distinction between two versions of the problem). This means that the fact of evil presents theism with a problem even though it does not allow us to deduce the falsity of theism. Granting that CSR and theism are or at least can be logically compatible, the former nevertheless has potential to contribute to the inductive problem of evil for theism. It does this by adding a puzzling evil to the stock of evils in the light of which theism is less probable. This evil consists in the immoral behaviour which results when humans succumb to cognitive susceptibilities that reflect the natural functioning of our minds. Why is the probability of theism negatively affected if we put this evil in the already existing basket of evils theism must account for? Why is this evil for example distinct from the evil of suffering brought about by natural disasters or by humans whose minds do not work normally? To answer this, we here follow a general line of argument Richard Swinburne (2004) pursues to show that that there is a good inductive argument from evil *against* theism.[75]

Swinburne suggests that to handle the existence of evil in general, theism must introduce heavenly compensation in an afterlife as an auxiliary hypothesis. This to some extent reduces the probability that theism is true, since the probability that a conjunction of two claims is true is lower than the probability that one claim is true. Swinburne also points out that, ceteris paribus, when it is necessary to add an extra explanation to a thesis, then that thesis becomes less probable. The upshot is this: it is less likely that God exists and that those who suffer receive compensation in heaven than that God exists. In logical terms, therefore, "P (theism| evil) < P (theism)" (Swinburne, 2004, 264-266). This is what it means that there is an inductive argument from evil. Now, let us by evil* mean the addition of the cognitive science of evil to the common notion of evil. With evil* we thus mean the facts of widespread natural and moral evils plus the fact that a large amount of the latter evils occurs as the result of the normal functioning of design-features of our minds. Is "P (theism | evil*) < P (theism | evil)? In that case, CSR, which

[75] Swinburne (2004) does not, however, argue that this argument tips the balance of the evidence in favour of atheism. To the contrary, Swinburne argues that in light of the broad range of factors he considers, it is more likely than not that theism is true.

provides the additional latter fact, contributes to strengthening the inductive problem of evil for theism. This seems to be the case. The reason is that an additional auxiliary hypothesis beyond the afterlife hypothesis seems needed to account for the additional element evil* introduces. For example, the notion of a heaven where those who suffer will be compensated does not provide a good explanation of why God by abstaining from intervention allows a large amount of natural and moral evils *and* actively designs our minds with a susceptibility to bring about a large portion of the latter evils. The problem is not merely about the magnitude of evil and that as this magnitude increases, the likelihood of theism decreases. Instead, the problem concerns the moral responsibility God has in virtue of creating a design whose *normal* functioning in normal environments increases that magnitude. This seems different from the responsibility God has in creating a world where an avalanche might suddenly occur, there is a risk of forest-fires during the summer, and where some anomalous individuals embark on a violent path. As for example Swinburne argues, God would allow such natural and moral evils to occur since their presence enables greater goods to obtain, such as compassion for those who suffer and heroism in the face of difficulty (Swinburne, 2004, 263-264, see also Taliaferro, 2004, 305-310). But the additional element in evil* requires that some further explanation, in addition to this sort of greater goods account or the afterlife-hypothesis, is offered. We need an account of why God is morally justified in setting up our minds in such a way that their normal functioning in normal environments leads, in many cases, to evil behaviour. We are not here denying that such an explanation is conceivable. The problem is something else: the composite whole becomes less probable as the number of auxiliary hypotheses we must add multiplies. Theodicy is less likely to succeed in light of evil* than in light of evil, since it is harder to show that God is morally justified in allowing a massive amount of evil brought about when we succumb to susceptibilities that he gave us, than to show that God is morally justified in allowing a massive amount of evil (of whatever origin). Since theodicy is less likely to succeed in light of evil* than evil, theism is less likely to be true given the former than the latter. The cognitive science of evil thus contributes to strengthening an inductive problem of evil.

Let us relate this conclusion to our overall discussion concerning normative implications. We have found normative implications for theist religion in sense (1): findings in CSR bear logically on the truth of theism in virtue of contributing to an inductive problem of evil (see Ch. 1, section 2). This finding is important given the topic of the present chapter, which is TE, a thesis connected to theism. Findings that bear logically on theism affect the overall balance of the evidence for theism and therefore have an impact on claim 2 in TE. This means that a normative implication in sense (2) for theist religion obtains. CSR is relevant in assessment of

an argument or thesis connected to theism.[76] Our findings also show that CSR has normative implications in sense (2) for theist religion in virtue of being relevant to the inductive problem of evil. This is because this argument represents a distinct example of an argument and thesis connected to theism, independently of its relevance to claim 2 in TE. Finally, the cognitive science of evil is relevant to theistic integration-theses, also independently of their potential relevance for claim 2 in TE. We have seen that such theses attempt to integrate central findings in CSR into a theistic framework. As mentioned earlier, Barrett (2004, see also Visala, 2011) hints at how such a thesis can be defended: Science provides information about the nuts and bolts of religion, and theist philosophy situates those nuts and bolts in a larger religious framework. It seems, for the parsimony-related reasons specified above, that the theist who finds CSR convincing and therefore embarks on an integration-attempt encounters a problem. To salvage an integration-thesis after we introduce the cognitive science of evil, this theist must propose a complicated theological explanation of why our minds would have a susceptibility to evil.[77] Given that this is necessary to salvage such a thesis, the science of evil creates the same problem as the one that we saw arising for theism itself. Theistic integration-theses become less probable as the number of auxiliary hypotheses needed to make sense of the scientific findings one attempts to integrate multiplies. But this does not mean that integration becomes impossible. For example, it is not necessary in such an enterprise to show that God is not morally responsible for the outcomes of natural human cognitive functioning. God could be morally responsible for allowing x yet morally justified in doing so. What seems necessary, and far from easy, is to show that God – who is perfectly good - is morally *justified* in giving us a cognitive susceptibility to evil. In conclusion, given how the cognitive science of evil in this way challenges general integration-attempts and given that these attempts are examples of distinct theistic theses in their own right (which may or may not be seen as relevant to claim 2 in TE), we have a further example of a normative implication in sense (2) for theist religion.

We now turn from a morality-concerned challenge to theism to a group of morality-based arguments *for* theism and look at what relevance CSR has for them. We thus further investigate the bearing such research has on TE.

[76] How serious is this implication? This depends on how strongly the probability of theism is affected, something I have not tried to assess. Our finding is nevertheless sufficient to establish a normative implication. It is, as specified in chapter 1, sufficient for normative implications in sense (2) that CSR is *relevant* in assessment of a theistic thesis or argument.

[77] Teehan (2016, 49-55, see also Teehan, 2018) considers theistic responses. The first concerns the Christian notion of original sin and the second involves the notion that human beings are spiritually and morally immature.

5. An evolutionary and cognitive perspective on moral arguments for the existence of God

A survey conducted by the Pew Research Center in 2013 found that globally, people widely believe that to be a moral person (presumably a morally good person) one must believe in God.[78] Support for this belief is strong in the Middle East, Africa, large parts of Latin America, and in the United States. In for example Egypt, as much as 95% was found to have this belief (Pew Research Center, 2014). At first blush, the psychological evidence seems to support this widespread idea. There is evidence that religious primes can induce pro-social (cooperative and honest) behaviour. For example, Brandon Randolph-Seng and Michael E. Nielsen (2007) found that subjects who were reminded of religious concepts when carrying out a seemingly unrelated task would subsequently cheat less than a control-group whose task involved non-religious concepts. As another example, Will Gervais and Ara Norenzayan propose that thoughts of God[79] can induce the feeling of being monitored, and that this can induce pro-social behaviour (Gervais and Norenzayan, 2012, see also Norenzayan, 2013, 33-47 for a survey of similar studies). Interestingly, the pro-social effects of feeling monitored by a *punishing* agency seems to be the strongest (Norenzayan, 2013, 43-47). Then, not any deity makes us good. Instead, as Norenzayan puts it, "mean gods make good people" (2013, 44).

These findings are intriguing, but do not plausibly show that one *must* be a religious believer to be a morally good person. For example, as Norenzayan points out, it is not the religious belief per se but brief religious reminders that induce pro-sociality (Norenzayan, 2016, 38-39). But more importantly, Norenzayan argues further that reminders of supernatural agents are not the only factors that induce pro-social behaviour. The feeling of being monitored and the reduced sense of anonymity this involves is the central factor, and this can be brought about by a variety of means (Norenzayan, 2013, 19-23). Conversely, factors that increase our sense of anonymity (such as wearing a mask or being in a dark-lit room), have the opposite effect: they decrease our willingness to engage in pro-social behaviour and lower our threshold to act anti-socially (Kurzban and Neuberg, 2005, 656, Zhong, Bohns, and Gino, 2010). Activating a religious notion is therefore not necessary to induce pro-social behaviour. Supposing pro-sociality has something to do with morality, a religious belief is therefore not required for being

[78] Not necessarily what I mean by "God" (as specified in Ch. 1, section 2, and further clarified in footnote 3).
[79] Again, not necessarily what I mean by "God".

morally good. Below, I point to a possible explanation of why many nevertheless link morality and religion together.

This brings us to what the current section is about, namely what relevance CSR has in assessment of a set of moral arguments for the existence of God[80] and how that relevance, if there is any, impacts on claim 2 in TE and P2 in I-ENT. The arguments we shall assess all involve a link between morality and religion but look at morality from different perspectives. The first argument concerns how to explain our ability to think and act morally. The second is about what can provide an adequate basis for something like *moral realism*. I understand moral realism as the thesis that there exist moral facts and properties[81] and that their existence or instantiation is constitutively independent of our opinions (Miller, 2013, 4).[82] Our third example is an argument that deals with both these issues.[83]

Helen De Cruz and Johan De Smedt (2015, 109-130) investigate moral arguments for the existence of God from the perspective of research in CSR. Their discussion is therefore a good place to start our investigation of whether such research is relevant in assessment of moral arguments.

The argument from moral awareness

Our first example is an argument De Cruz and De Smedt (2015) call "the argument from moral awareness", attributed to Richard Swinburne (2004), John Hare (2004), and C.S. Lewis (1952):

P1: To explain the moral sense, we can invoke either theistic or naturalistic explanations.

P2: Naturalistic explanations cannot account for the moral sense.

P3: Theism can account for the moral sense.

C: Therefore, the best explanation for the moral sense is theism (De Cruz and De Smedt, 2015, 111).

[80] "God" in the sense specified in chapter 1, section 2. That is, roughly something like the God of Abrahamic theism. I treat the 3 arguments we shall here discuss as arguments for the existence of God in that sense.

[81] An example of a moral fact, if there are any, is that killing an innocent person is wrong. In terms of moral properties, a moral realist can say that such an act has the property of being wrong.

[82] Moral realism is a *cognitivistic thesis.* We can understand cognitivism as the claim that moral judgements express beliefs and that these are truth-apt, meaning either true or false (Miller, 2013, 3). Moral realism, as specified here, involves a metaphysical thesis about moral facts and properties and an objectivist thesis about how those facts and properties are independent of our opinions about them. Below, we shall consider an argument for theism called "the argument from moral objectivism". In discussions of this argument (see for example De Cruz and De Smedt, 2015, 119-120), the term "moral objectivism" is applied for something like moral realism. For simplicity, I use "moral realism", which seems to be the most common metaethical term.

[83] There are also moral arguments from human dignity (Linville, 2009) and practical moral arguments. The present investigation does not discuss these. For an extensive overview of moral arguments in general, see Evans (2018).

In this argument, the moral sense can be viewed as a tendency to form spontaneous, evaluative beliefs about the character and behaviour of others (De Cruz and De Smedt, 2015, 110). As De Cruz and De Smedt (2015, 112) point out, those who advocate this argument see it as in principle impossible for naturalistic explanations of the moral sense to succeed. This indicates that P2 is a quite strong claim. P3 seems to be the claim that theism in principle can account for the moral sense. The door is then open for the theist to present a specific account of the formation of this sense. We assume this interpretation of P3. How does research in CSR bear on this argument?

The relevance of naturalistic research on morality is quite notable when it comes to P2: this research indicates that the moral sense *can* be explained naturalistically. To start with, there are good Darwinian explanations of the tendency in human and a variety of non-human animals to engage in altruistic behaviour, that is, behaviour whereby the organism incurs costs to its individual fitness by acting in ways that increase the fitness of another organism. The primary mechanisms that explain these behaviours are kin-selection and reciprocal altruism (Boyd & Silk, 2015, 173-191, see also Teehan, 2010, 21-27). A third mechanism, indirect reciprocity, can explain our predisposition to act altruistically towards non-related individuals who we might not interact with again or who might not be likely to reciprocate directly. The central principle here is that by acting altruistically under those circumstances, the subject obtains a better social reputation and is for this reason more likely to benefit from cooperation with others in the future (Teehan, 2010, 27-32, De Cruz and De Smedt, 2015, 113-114). John Teehan (2010) proposes that we see these mechanisms as basic "layers" that make up an evolved morality. Thus far then, it seems that science can explain a range of pro-social behavioural proclivities. Can it also explain the moral sense? As Teehan further points out, there is a growing consensus in the scientific community that we have a predisposition to form intuitive, non-reasoned, emotionally laden moral judgements in a large variety of social situations, and that after forming such judgements, we tend to form "post hoc rationalizations" by which we justify those initial judgements (Teehan, 2010, 16-20 see also Krebs, 2005, 766). This seems a plausible construal of the moral sense. What is its (ultimate) origin? This disposition to spontaneously form evaluative judgements has according to numerous scholars evolved as a solution to recurrent adaptive challenges relating to the need to interact socially with conspecifics, carry out complex coordinated tasks, and to form and maintain groups (Tomasello, 2016, Teehan, 2010, Krebs, 2005, De Smedt and De Cruz, 2015, 118, Joyce, 2006, Boyer, 2002, 204-205). In light of this, naturalistic science can explain how the human moral sense arose. This does not mean that science currently has a complete understanding of this complex issue, but it strongly illustrates

the viability of naturalistical explanation of morality. This strongly undermines P2 – which states that it is in principle *not* possible to explain morality naturalistically. In light of this, CSR has negative relevance for this argument by undermining this central premise.

Do we have grounds for establishing any further relevance? This brings us to P3. It does not follow from what we have seen thus far in this section that P3 is wrong. The viability of naturalistic explanations of the moral sense does not in itself rule out the general prospects for a theistic alternative. It could for example be that both theistic and naturalistic explanations can, at least in principle, account for the moral disposition. De Cruz and De Smedt, who also reject P2 in the argument above (2015, 119), point to the possibility of a weaker version of the argument, which recognizes that possibility:

> "It might be more prudent to consider a weaker version of the argument from moral awareness, where theism is regarded as the best explanation for human moral intuitions, without, in principle, excluding naturalistic explanations." (De Cruz and De Smedt, 2015, 112).

This version of the argument is not ruled out by the mere ability of CSR to explain morality. Assessing this version of the argument would therefore require more than illustrating that ability. A comparative analysis that establishes the relative strengths and weaknesses of each type of account of morality and correctly judges whether theism provides the best option seems needed. I shall not pursue that here but instead point to a challenge I suspect will create problems for theistic accounts. In light of "the cognitive science of evil" we discussed in the foregoing section, theistic accounts of morality face a distinct explanatory challenge. For example, if divine intervention or guidance is involved in the formation of human morality, then our cognitive susceptibility to evil, which is a normal part of that morality rather than an example of its perversion, becomes a puzzling fact that requires an explanation. Why would a loving God design our moral thought with the in-group and out-group bias current studies reveal? We shall return to this issue below when discussing two other versions of the moral argument. As a tentative conclusion, the cognitive science of evil issues a challenge that weaker versions of the moral argument from awareness must deal with. This challenge extends to P3 in the strong version of the argument. But that does not mean it provides sufficient grounds to reject P3. That is, it does not follow from the thesis that our morality comprises a susceptibility to evil that theism in principle cannot succeed in accounting for morality. Instead, this thesis encumbers theism with additional explanatory burdens. To conclude, CSR bears negatively on the strong version of the moral argument from awareness by strongly undermining P2 and by

creating a challenge for P3. The naturalistic explanation of morality CSR provides therefore strongly undermines this argument. We have found a normative implication for theist religion in sense (2): CSR is relevant in assessing an argument or thesis connected to theism.

The argument from objectivity

De Cruz and De Smedt also consider a second argument from morality, which they call the argument from moral objectivism. In this argument, attributed in slightly different versions to Robert Adams (1987) and Paul Copan (2008), moral objectivism is the view that morality is completely independent of human beliefs, conventions, social institutions, and so on, and instead depends entirely on non-natural properties such as moral facts (De Cruz and De Smedt, 2015, 119-120). It thus seems an example of moral realism. Paul Copan's formulation (quoted in De Cruz and De Smedt, 2015) goes as follows:

P1: If objective binding moral facts exist, then God exists.
P2: Objective binding moral facts exist.
C: Therefore, God exists (De Cruz and De Smedt, 2015, 120).

Since these premises do not make claims about human evolution or psychology but metaphysical issues, it is less clear than in the previous example that empirical science can help us appraise their truth, at least directly. One way in which psychological studies nevertheless can be relevant concerns the extent to which the premises can be considered something like common sense or in general intuitively true. As De Cruz and De Smedt (2015, 121) point out, some supporters of the moral argument take the (alleged) common sense nature of moral realism as prima facie evidence for its truth. One might also contend that since moral realism has this status, those who deny rather than those who endorse this position have the burden of proof. Then, a psychological study of whether and to what extent moral realism is common sense or strongly intuitive is relevant at least to the plausibility of their defence of this position. A psychological study of whether the claim P1 states is common sense or in general intuitive seems relevant for similar reasons.

Let us start by looking at P1. This premise suggests that morality is objective only if God exists. This relates loosely to the idea famously discussed in *The Brothers Karamazov* by Dostoyevsky that if there is no God, then everything is permitted. A way to construe this

premise is as the claim that *only* the existence of God can provide a basis for moral realism. Let us proceed to discuss the argument given that interpretation.

Two quite different considerations, the second of which involves research in CSR, have a bearing on this premise. The first seems the most serious, but we shall not focus on it in the present investigation, since it primarily concerns metaethics rather than psychology. Two points will illustrate. First, there are many current defences of moral realism from a non-theistic perspective. As examples of this, consider Sturgeon (1988), Boyd (1988), Brink (1989), Railton (2003), and Shafer-Landau (2003).[84] If any of these succeed in establishing a realist theory, then morality has been shown to be objective without the need for an appeal to God. In that case, P1 can be rejected. Second, theistic metaethics faces distinct challenges of its own. In particular, theistic metaethics based on so called "divine command theory" faces a version of the famous Euthyphro dilemma.[85] Unless this problem is solvable by the theistic metaethicist, P1 is again in trouble.

The second relevant consideration bearing on P1 concerns psychological studies that shed light on how subjects intuitively relate morality and religion. Let us ask how well spontaneous moral intuitions map on to what P1 states.

---

[84] For an overview of current non-theistic defences of moral realism, see Miller (2013), 143-276.

[85] Divine command theory holds that moral properties such as rightness, wrongness, goodness, and badness are grounded in what God wills or commands (Berg, 1997, 525). The dilemma such a theory faces is this: Does God decide that X is good because X is good or is X good because God decides that it is good? On the first horn of this dilemma, the challenge is to show that the objectivity of morality depends in a significant way on God. If God needs to consult an independent moral standard when making moral judgements, then that does not seem to be the case. Furthermore, we then need an (independent) account of the moral standard God consults. On the second horn, morality depends on God because God decides what is right and wrong and grounds the objectivity of morality in his commands or in his will. A challenge then arises. We need an answer to the following kind of question: If the goodness or wrongness of acts depends entirely on what God commands (or wills), is murder or rape good if God commands or wills these acts to be good? (Berg, 1997, 527-529) Adams (1987) defends a modified version of divine command theory which he believes can handle this challenge (see also see Berg, 1997 and Evans, 2018 for an account of this dilemma and various theistic responses to it). Related to the latter horn is an epistemological issue it is worth mentioning. How can we know what God commands or wills? Those who appeal to the Bible for an answer face an interpretational challenge. Presumably, God does not view warfare and murder as good, but this text depicts God otherwise. For example, God is portrayed as commanding brutal invasions, and even the murder of women and children. Consider for example Deuteronomy 20: 10-20. Biblical scholars, obviously aware of the challenges such texts pose, have for more than two millennia dealt with these issues, and devised a range of interpretational methods and perspectives, many of which do not involve a literal reading of the Bible (see Grant and Tracy, 2005 for a historical overview). This bewildering multiplicity of methods employed testifies to the difficulty of obtaining reliable information about God's commands or will from this text. Interestingly, research in CSR is relevant to Biblical interpretation and it seems to add to this difficulty. As John Teehan (2010, 2016b) argues, the Bible, including the New Testament, gives expression to an underlying evolved moral psychology. For example, this text shows evidence of being afflicted with in-group and out-group bias. Consider for example Teehan's evolutionary interpretation of the Decalogue (2010, 80-101), the Gospels, and the letters of Paul (pp. 104-143) to appreciate this. To the extent that such a bias afflicts the Bible as a whole, rather than only a minor selection of passages, and given that this text purportedly can tell us something significant about God's commands or will (despite this), the interpretational challenge for the theistic metaethicist seems serious indeed.

As a start, human moral intuitions in general and what P1 states are not about the same things. The latter is about how moral realism and God are linked together. By contrast, evolved evaluative intuitions are about aspects of hominid social interaction, especially those that pertain to group-cohesiveness and the danger of free-riding (Joyce, 2006, Tomasello, 2016). What P1 states is therefore not intuitive merely in virtue of the natural functioning of the mind, the way for example naïve physics or naïve biology is (see Ch. 2, section 4). However, given the right cultural input, something like P1 is likely to seem true to many subjects. This relates to the widespread belief about religion and morality illustrated earlier. Let us apply the account of the theistic disposition (Ch. 3) and explore how. Based on that account, it seems that with the right cultural exposure, what P1 states *can to some extent become intuitive*. How does that happen? As a start, general cognitive biases (see Ch. 2) we all share make us susceptible to acquire and entertain MCI concepts, and a specific variant of such concepts that depicts *morally concerned* deities is particularly cognitively attractive (Boyer, 2002). This variant is attractive since these deities are intuitively conceptualized as having full access to "socially strategic information", and our evolved mechanisms for handling various social tasks make us especially sensitive to such information (Boyer, 2002, see also De Cruz and De Smedt, 2015, 122-123). Cultural exposure to the concept of God introduces a cognitively attractive religious notion to minds already evolved to care about morality (in the sense of having the moral sense or disposition discussed earlier). As I have argued (Ch. 3), such a concept will gradually increase its relevance by becoming activated in concrete situations where evolved systems trigger intuitions on which the concept then latches. This includes systems for moral thought. In light of this, such entanglements can over time contribute to creating an intuitive link between God and morality in the mind. With this link established, the link between God and morality in P1 might begin to seem intuitively true. Given the need for the right cultural exposure for this to happen, De Cruz and De Smedt (2015) seem right when stating that there is something culturally specific about the supposed link between morality and God in moral arguments:

> "The intuitive connection between theism and morality that forms the basis of the moral argument may thus find its origin in our cultural history: Western theists (and non-theists) have been raised in large-scale societies that historically (or still) hold belief in a high god." (De Cruz and De Smedt, 2015, 124).[86]

[86] According to De Cruz and De Smedt (2015, 124-125) this does not fully explain how theism and morality are linked. There is also evidence, they point out, that subjects in non-theistic cultures intuitively link supernatural agents (like ancestor-spirits) to morality: The Tyvans of Mongolia implicitly seem to believe that their "spirit masters" have an interest in morality (Purzycki 2013). However, there is also a difference: Tyvans do not, as do

Another way to appreciate the cultural specificity of the above link is this: Without the right cultural input which helps make God and morality seem related, morality will be largely the same. As Pascal Boyer for example puts it, we do not have morality because we have religion. Instead, human evolved morality makes religious concepts easier to acquire (Boyer, 2002, 193, see also Pyysiäinen and Hauser, 2009). The moral intuitions religious concepts become relevant by parasitizing would be the same anyway. Recall for example (Ch. 3, section 3) how S's belief that God resents Z just is the output of S's cheater-detector but with a theistic gloss added. Without that gloss, S would simply resent Z. This indicates that we intuitively believe that behaviours are good or bad irrespective of whether we think of them in terms of what God or any other deity thinks about it. Without God or other supernatural agents latching on to our thought, cheating and other forms of anti-social behaviour corrosive to the internal stability of groups intuitively seem wrong. Then, our intuitive moral thoughts independently of specific cultural inputs will not on their own bring us to accept what the character in the *Brothers Karamazov* states, that unless there is a God, "everything is permitted". A wide range of behaviours are intuitively viewed as "not permitted" whether or not one entertains the notion of God. But as Boyer (2002) has shown, given the introduction of the concept of God to the mind, what is permitted or prohibited might be reconceptualized as what God permits or prohibits.

How does this account bear on P1 in the objectivist argument? We have undermined the notion that this thesis in general is a common sense or intuitively compelling one, given the conclusion that cultural input is necessary. P1 is probably not intuitively compelling across a range of representative cultural divides, but it can become intuitive in some contexts. This creates trouble for those, if there are any, who appeal to the intuitiveness of P1 as prima facie evidence for its truth. But our account does not tell us that P1 is false. The soundness of philosophical arguments for and against P1, which is a substantive metaethical thesis, seems the most decisive factor. I pointed to defences of non-theistic moral realism and challenges to theistic metaethics largely (but not entirely: see footnote 85) unrelated to CSR as the most serious challenges to P1. So, in conclusion, CSR does not strongly impact the first premise of the objectivist argument to morality.

Let us now turn to P2 in the objectivist argument, the thesis that there are moral facts and properties and that these are independent of human opinions about them. This thesis is an

for example Christians, believe that their spirits will punish moral transgressions (De Cruz and De Smedt, 2015, 125, cf. further discussion in Teehan 2016b).

example of moral realism. I do not directly ask the substantive metaethical question whether this thesis is correct, but instead look at what research in CSR can tell us about whether we are intuitive moral realists. I shall then briefly deal with what relevance, if any, such research on intuitive moral thought can have in assessing moral realism (i.e. the metaethical position). This allows us to expand on the picture of moral thought presented thus far (in the current and previous section). We shall briefly look proximately at the development of morality, the phenomenology of forming evaluative thoughts, and then appreciate some of the patterns in such thought. In this way, we can find out if intuitive moral thought corresponds with what P2 in the objectivist argument states.

As a start, there is evidence that already early in childhood, intuitive thought about morality and about social conventions are distinct. For example, subjects readily distinguish between on the one hand actions such as harming someone or distributing goods unfairly, and on the other, failing to comply with a social convention (concerning for example how to dress or behave at the dinner table) (Turiel, 2008, 491-494). Moreover, as Turiel points out, there is evidence that children tend to focus on the consequences of harmful or unfair treatment rather than what an authority figure says about it when forming judgements about its wrongness (Turiel, 2008, 491, see also De Cruz and De Smedt, 2015, 121-122, Boyer, 2002, 203-205). Moreover, and this relates more directly to the question of whether we are intuitive moral realists, studies support the notion that children and adolescents view moral issues concerning welfare, harm, fairness, and so on, as obligatory and not dependent on what social and cultural conventions indicate (Turiel, 2008, 493, cf. Buchanan and Powell, 2018, 135-137). So, intuitive moral judgements are made about actions being right or wrong or about distributions of goods being just or unjust independently of what we are told in our social and cultural context. This aspect of such judgements seems related to the objectivist claim in moral realism, which is that moral facts and properties are independent of our beliefs about them. Here our intuitive moral thought seems roughly in line with realism, and this can help explain why P2 will seem intuitively true. Moreover, in contrast to the intuitiveness of P1, this intuitiveness does not appear to hinge on having the right cultural exposure. Instead, depending on the cross-cultural generalizability of the relevant studies, this intuitiveness seems to hinge more on cognitive biases and tendencies we all share.

Another relevant point concerns the phenomenology of forming moral judgements: they seem objectively true rather than as mere opinions. As for example Boyer (2002, 205) contends, our intuitive moral judgements come with a “realistic bias” in the sense that we intuitively believe that they represent things the way they actually are. For example, when forming the

judgement that stealing is wrong, we do not think of this just as our subjective opinion or as prescribed by a social convention. It seems objectively true (Boyer, 2002, 205). The difference in the phenomenology of holding subjective opinions ("Coke is better than Pepsi") and moral beliefs ("murder is wrong") can be thought of in terms of what Stephen Darwall calls "objective purport". In terms of how having moral beliefs is experienced by the subject, moral beliefs purport to truth in a way that subjective opinions do not (Darwall, 1998, 18-19). The realistic bias or inbuilt purport to truth described here seems related to moral realism, which comprises an objectivist thesis. From here we can envisage an additional step to the metaphysical thesis it also comprises (i.e. the thesis about moral facts). Intuitive explanatory tendencies can help us take that step. For example, subjects might intuitively explain the realistic bias by postulating that there are moral facts they get right in virtue of forming moral beliefs. However, there are also findings in cognitive science which point in a very different direction and indicate that we are *not* intuitive moral realists. These relate to the cognitive science of evil discussed in the foregoing section and some far from benign patterns in the content of our moral thought.

Think again of the objectivity moral realism sees moral facts and properties as having vis a vis human opinions, conventions, and practices. For the moral realist, a moral fact completely independent of facts about something as petty as human group membership or ethnicity determines the moral status of an action like hurting a defenceless person for no reason. Such an action is wrong independently of what ethnic groups the victim and perpetrator belong to. But as we saw in the previous section, human moral psychology involves strong in-group favouritism. This might lead bystanders belonging to the same group as the perpetrator to spontaneously see the violence as justified, or less morally wrong than they would have if the victim were a member of their own group and the perpetrator an out-group member. Moreover, they are less likely to empathize if the victim is an out-group member (see for example Boyer, 2018, 38-39, Tomasello, 2016, 91-92, Buchanan and Powell, 2018, 130-132).[87] These biased judgements and reactions which (we may assume) do not correspond with moral facts (if there are any) will to the subjects nevertheless seem entirely correct. That is, there will be the same realist bias and purport to truth at play as when one (in another context) forms the belief that "murder is wrong". Xenophobic or hateful beliefs about an out-group member of a different ethnic origin might in a context where strong out-group bias is triggered seem just as

[87] We should be aware that Buchanan and Powell argue that even though our moral psychology contains such a feature, that does not mean we are "hard-wired" to act in tribalistic ways. Rather, depending on what sort of environment our moral psychology is expressed in and how it develops as a result, the outcome could be either tribalistic behaviour or more morally inclusive thought and behaviour (see Buchanan and Powell, 2018, 135).

objectively true as intuitive moral beliefs of a benign nature formed in another context. That these biased judgements seem true rather than akin to mere opinions (about Coke or Pepsi for example) achieves something important. As Teehan points out, being strongly convinced is a particularly effective way to convince others and to signal one's commitments (Teehan, 2010, 192-193, see also Boyer, 2018, 49-52). In light of this, our intuitive moral thought departs from the moral realist view that facts and properties independent of our beliefs about them determine the truth of our moral beliefs. Our thought departs from this view in virtue of allowing petty, coalitionally afflicted concerns and ethnic priorities which presumably are irrelevant with respect to the moral rightness or wrongness of actions to affect our evaluations of rightness and wrongness.

The above findings and contentions concerning our intuitive moral thought are complex. We saw that some aspects of our moral thought seem related to moral realism. However, in light of how strongly an in-group and out-group bias afflicts such thought, we are not plausibly depicted as intuitive moral realists. Then, the basis for the prima facie intuitive plausibility of P2 appears weakened, something that bears negatively on attempts to defend it based on its alleged intuitiveness. But the failure of our intuitive moral thought to map neatly on to what P2 states does not in itself indicate that P2 is false. It could also be taken to indicate that our intuitive moral thought is liable to get moral facts wrong, at least in contexts where in-group/out-group bias is triggered. But that does not entail that there are no objective moral facts. Furthermore, P2 is after all (as is P1) a substantive metaethical position defended by a range of current theories (indicated earlier). To appraise the truth of P2, the merits of metaethical arguments for and against moral realism seem more decisive than the content of our intuitive moral thought. Before concluding, let me briefly flag another strategy by which one can try to show something stronger than what I have done.

As has been suggested, one could interpret Michael Ruse and E. O. Wilson (1986) as arguing that evolutionary considerations about the nature of morality indicate that morality is an illusion (see De Cruz and De Smedt, 2015, 121-122, and Linville, 2009, 398). Ruse and Wilson for example argue that our genes "deceive" us to think about right and wrong as reflecting objective properties that are binding on us. Furthermore, in this way, the "illusion of objective morality" is created in us (Ruse and Wilson, 1986, 179-180). A consideration relevant to this sort of evolutionary debunking argument is this: evolutionary science offers explanations of how moral convictions arise and why they seem objectively true without needing to postulate the existence of moral facts and properties (Teehan, 2010, 186-193, Joyce, 2006, 183). One might then be tempted to suggest that this explanatory redundancy provides reason to *deny* that

there are moral facts and properties altogether. De Cruz and De Smedt argue that this is not plausible. The evolution of morality does not show that morality is an illusion, since evolved moral capacities could potentially track moral truths (De Cruz and De Smedt, 2015, 122). Another issue for this sort of argument concerns the danger that it commits a fallacy. As I stated (Ch. 1, section 5) in regard to J. Barrett's construal of debunking-arguments against religious beliefs, it seems fallacious to argue that claims or beliefs are false merely on the basis of a theory of how we come to accept or form them (see also Kahane, 2011, 105). This also applies to P2. Finally, as mentioned above, there are many current defences of moral realism, and their merits seem the most decisive in settling the dispute concerning the objectivity of morality. It seems that the primary factor that would make moral realism true is the existence of moral facts and properties that obtain independently of our beliefs about them. What then would make this position false is the non-existence of such things. This indicates that a psychological study of moral intuitions related to this metaethical position does not on its own establish its truth-value. This seems so whether we are intuitive realists or, as I argued, our moral intuitions depart from moral realism in some respects. For this reason, this debunking-strategy seems to go too far.[88]

In conclusion, the main factors bearing on the truth of the premises in the objectivist argument involve metaethical issues largely unrelated to CSR. Research in CSR can nevertheless to some extent be relevant in virtue of challenging attempts to defend it which appeal to moral intuitions, and it is relevant to those interested in whether moral intuitions map on to what the premises state. This means we have in this case not found strong (negative or positive) relevance for a theistic argument. Let us turn to our third and final example of a moral argument.

Swinburne's moral argument

Richard Swinburne presents a moral argument as part of his larger cumulative inductive argument for the existence of God (2004). It is a hybrid that combines elements of the two

[88] This, to be more specific, is my tentative conclusion in the present investigation. We should be aware that there is an ongoing debate in current metaethics concerning the implications evolutionary considerations have for moral realism. For example, Sharon Street (2006) applies such considerations to show that moral realism faces a serious dilemma in the light of which we should opt for moral antirealism. For critical discussions of this argument, see for example Shafer-Landau (2012) and Fitzpatrick (2014). We should also be aware that this (tentative) conclusion does not necessarily rule out the viability of debunking in a different sense. We can distinguish attempts to use evolution to show that *moral realism is wrong* and attempts to use evolution to *undermine the justification of moral beliefs* (Kahane, 2011). Richard Joyce, for example, attempts the latter. He argues that evolutionary considerations undermine the epistemic status of moral beliefs in virtue of explaining such beliefs in ways that make moral facts redundant (Joyce, 2006, 179-184).

former arguments. By "morality", Swinburne means an ability to form evaluative beliefs. This ability involves the employment of moral concepts in evaluations of the moral qualities of actions. Furthermore, with this ability we can make moral evaluations independently of what our immediate intuitions and desires tell us. We can, for example, realize that some of our intuitions or desires are bad, or that something we do not spontaneously want to do nevertheless is the right thing to do. Secondly, by having this moral ability, we can become aware of *moral truths*. For example, if one judges of a morally good action that it is good and of a morally bad action that it is bad, one becomes aware of such truths (Swinburne, 2004, 215-216). This is a hybrid conception of morality, which includes something like the notion of a moral sense discussed earlier as well as a moral realist thesis according to which the truth of moral beliefs consists in a correspondence with a moral fact or property. Swinburne does not argue to theism from the moral disposition or from moral objectivism in isolation, but he argues from the thesis that we have this disposition and that when using it we can obtain moral knowledge. According to Swinburne (2004), the existence of morality (in this hybrid sense) provides evidence for the existence of God because God is what best explains it.

To show this, Swinburne argues as follows. While neo-Darwinian science provides explanations of phenomena like altruism and pro-sociality, unless we presuppose the existence of God, the existence of the ability to form moral *beliefs* like those specified above is a puzzle. That is, we can tell a story about how a mechanism like natural selection could build the basis for pro-social behaviours and altruism, but evolutionary theory provides no good reason to expect the presence in us of tendencies to form moral *beliefs*. Furthermore, such a naturalistic theory does not account for how there can be moral truths that we become aware of. We must, according to Swinburne, appeal to theism to explain this. God has reasons to endow us with such a tendency because this helps ensure our ability to make moral choices, and this is something God wants. God could ensure this, Swinburne suggests, by connecting to our genes a causal belief-producing program that leads us to form moral beliefs (Swinburne, 2004, 216-217).

Given the strong claims about the inability of science to explain evaluative belief-formation, CSR-research bears strongly on this argument. Let us start by highlighting that it is problematic to bring pro-sociality and altruism apart from our evaluative belief-forming capacity, and to suggest that the latter is a mysterious puzzle for the naturalist. In our species, evaluative thought is integral to the two former phenomena. The moral sense (discussed in relation to the first moral argument) is a disposition to form evaluative *belief*, not simply a mechanism for inducing pro-social or altruistic behaviour (for evidence of this, see the

discussion of the first moral argument). Recall also the example (Ch. 3) concerning S's negative evaluative judgements about Z's cheating. The formation of such beliefs is not an optional add-on for the mechanism that governs their production. It is part of its operation (Cosmides and Tooby, 2005) Therefore, the fact that we can form evaluative beliefs is not a puzzle for the naturalist.

The fact that we can also take stock of our intuitive morality and form beliefs about its rightness and wrongness is also no puzzle. As John Teehan (2010) argues, evolutionary theory can account for the emergence of the capacity for intuitive, evaluative thought. Our intuitive, evaluative thought strongly reflects the kind of emotional reactions we have to different types of social encounters or situations. Reflective thought whereby we take stock of such thought and form further beliefs about it often functions to provide us with post hoc rationalizations of our immediate intuitive thought. Both are in a larger sense part of human moral psychology and not something science cannot explain (Teehan, 2010, 9-42).

Swinburne might here object that there is still a puzzle when it comes to the phenomenon that we form moral beliefs according to which something we desire is bad or wrong or according to which something we do not desire is good or right. If our reflective moral beliefs simply rationalize more immediate emotionally charged intuitions and reactions, then one would expect the former not to judge the latter in a negative way (but instead justify it). Richard Joyce's account of morality contains an answer to this. Joyce is interested in the possible fitness-benefits of having a conscience, that is, having the capacity to judge oneself morally and to be capable of remorse and guilt. Joyce asks why it would be adaptive to form moral beliefs about one's own conduct or one's own behavioural inclinations and tendencies (Joyce, 2006, 108). The answer concerns the motivating and behaviour-regulating effects of forming moral beliefs. Forming such beliefs can motivate us to do what we judge to be right and to avoid doing what we judge to be wrong even though our immediate desire or inclinations nudge us in the opposite direction. For example, by morally judging that cheating during social exchange is wrong, one becomes committed not to cheat and this commitment stops one from further deliberating about the situation in terms of cheating. This can help the subject avoid acting on an immediate desire for short-term gain that in the long run is more fitness-costly than a more cooperative strategy (Joyce, 2006, 109-114, cf. Tomasello, 2016, 85-128, and the discussion of supernatural punishment theory in Ch. 2, section 2). In light of this, even moral beliefs whereby one judges one's own immediate inclinations to be morally bad is no puzzle for the naturalist.

A puzzle emerges, by contrast, if we opt for the theistic solution Swinburne proposes, which involves the notion of a divinely implanted, moral belief-inducing program. This is the

mechanism Swinburne introduces, presumably to illustrate how a theistic explanation of morality could work on a proximate level. But given the introduction of this mechanism, we need an answer to the question why our evaluative thought, if produced by a specific program God deliberately designed, is afflicted with the coalitional biases discussed earlier. The explanatory challenge that the cognitive science of evil creates haunts this proposal. The problem is not that these challenges cannot be handled from a theistic perspective but that in handling them the resulting account becomes complicated and less likely to be true. This reduces the prospects of such an account of being the one which *best* explains morality, something it must if morality is to provide evidence for God.

There is one part of Swinburne's conception of morality, however, which theism in fact appears to handle better than naturalistic empirical science. This is the ability of the moral disposition to enable awareness of *moral truths.* Naturalistic empirical accounts of morality do not readily explain this. But that is because they are dealing with something else than metaphysics (moral facts and properties) and epistemology (our moral awareness or knowledge) in the first place. It is therefore not surprising that such accounts appear less capable of handling morality in this sense than theism, a philosophical claim. However, that they are less capable than theism in achieving something they never aimed at in the first place, does not tell us that theistic accounts are the *best* accounts. We should not pit Darwinian empirical science against theism as if these were the two main alternatives. That is a false dichotomy. A fair and balanced treatment of the best non-theistic and theistic *philosophical* accounts of morality which shows that the latter are the best is required. As mentioned earlier, there is a wide range of non-theistic philosophical accounts of moral realism which attempt to deal with both the metaphysical and epistemological questions. Given that Swinburne's (2004) argument fails to consider such accounts, which seem the most promising alternatives to theism, his argument does not establish more than that theism gives us a possible explanation of morality. This does not suffice to show that morality provides evidence for the existence of God.

In this case, we have found that CSR is relevant in assessment of an argument for theism. The relevance in this case is negative and it seems more direct than the relevance established for the previous argument. This is because part of Swinburne's conception of morality involves a thesis about the inability of naturalistic science to explain the formation of moral beliefs. In light of evolutionary and cognitive research on morality we have grounds to deny this, and it is sufficient for this denial that such research at least in principle can explain such beliefs. Furthermore, by introducing a specific, theistic mechanism as an illustration of how moral beliefs arise, Swinburne's thesis is also challenged by the cognitive science of evil. By contrast,

the more philosophical part of Swinburne's conception of the moral sense, which involves a thesis about moral facts and how we can know them, is not affected directly by CSR. In regard to that part, the main problem seems instead to be failure to engage with the most appropriate non-theistic alternatives, namely non-theistic *philosophical* accounts of moral realism found in current metaethics.

In concluding this section, CSR has relevance in assessment of three moral arguments for the existence of God, the argument from moral awareness, the argument from objectivism, and Swinburne's hybrid moral argument. This relevance is negative with respect to the first and the last, and of limited strength with respect to the objectivist argument. Together with the implications of the cognitive science of evil, this result indicates that CSR can affect the balance of the evidence for theism, and therefore claim 2 in TE (the claim that the balance of the evidence favours the truth of theism). We shall not try to assess the extent to which this result impacts on that claim. The balance of the overall evidence for theism hinges also on a large variety of issues besides those we have considered. Nevertheless, given that there is impact, TE is affected. To relate this to our overall aim in the present part of the investigation, we have here found that CSR has normative implications in sense (2) for theist religion. This is because TE is a thesis connected to theism and CSR has been found to have relevance in assessing it. Moreover, the relevance for each of the moral arguments we have considered establishes distinct normative implications in sense (2) as well, since these are theistic arguments in their own right. Finally, our findings also bear on the internalist argument to neutrality, I-ENT by providing strong grounds to reject P2. Recall, that premises states that CSR does not shed light on *any* internal factor, such as reasons, arguments, or evidence. In light of what we have seen, this is not correct. When combined with the criticism of P1 offered in section 3, we now have strong grounds to reject I-ENT.

We now turn to another central position in the epistemology of religion, namely reformed epistemology.

Chapter 5. Reformed epistemology and the notion of a god-faculty

1. Introduction and chapter-outline

Consider the prototypical exchange between theistic evidentialists and evidentialist critics of religion mentioned in chapter 4, section 2. In this exchange, the protagonists agree that the justification of belief in God hinges on whether there is evidence for his existence. Their disagreement is about whether the evidence justifies the belief in God. Given the mentalist internalism we discussed in chapter 4, section 3, it is only possible to obtain well-founded (doxastically justified) beliefs about God by basing them on good evidence. But what if it is not necessary for the theist to base beliefs on evidence to form well-founded belief? What if the reliability of belief-forming processes or facts about the environment also can contribute to confer justification? This is a possibility externalism plausibly allows. We will now discuss an externalist thesis in the epistemology of religion, which if viable, allows us to bypass the above type of exchange. This is reformed epistemology (RE). Alvin Plantinga (2000), who together with Nicholas Wolterstorff is the central architect of this theory, denies that to have what he calls *warranted* theistic beliefs one *must* base them on evidence. He thus disagrees with both the theistic evidentialist and the evidentialist critic about evidence being necessary. We are here, as is clear from my discussion of evidentialism in the previous chapter, in agreement. In the next chapter, we will see how the specific externalist view I favour nevertheless departs from Plantinga's theory. Given that we now are in a general externalist landscape, we can turn our focus directly to Plantinga's substantive claims about the epistemology of *theistic beliefs*. The central claim is in short that *if God exists*, then the proper functioning of our minds confers warrant on theistic beliefs. A central notion related to this claim is that if God exists, then we have a *god-faculty*, sometimes also called *Sensus Divinitatis*. It is the functioning of this system that confers warrant.

What can cognitive science tell us about this? Kelly James Clark and Justin Barrett (2010, 2011) have brought RE together with research in CSR, contending that there now is empirical evidence for something like a god-faculty, and that this shows that RE and CSR might

describe something similar from different perspectives. According to Clark and Barrett, RE and CSR converge in their views of how forming religious beliefs is cognitively natural:

> "Reformed epistemology holds that belief in *God* is basic – that is, belief in God is a natural, non-inferential belief that is immediately produced by a cognitive faculty. Cognitive science of religion also holds that belief in *gods* is (often) non-reflectively and instinctively produced – that is, non-inferentially and automatically produced by a cognitive faculty or system." (Clark and Barrett, 2010, 174, my emphasis)

This is an intriguing attempt to bring cognitive science to bear in a discussion of religious epistemology. As we shall see, Clark and Barrett construe the god-faculty differently from how Plantinga does this. We can distinguish a Plantingian version of the god-faculty thesis and a version based on the interpretation Clark and Barrett defend. The distinction between "God" and "gods" relates importantly to these. I believe it points us to a way in which RE and CSR *diverge* from each other, namely in virtue of having different views of what sort of beliefs that are cognitively natural (in a way we shall discuss below). These views have different implications for the epistemic status of theistic beliefs. In RE, "God" means, roughly, the deity of Abrahamic theism, which Alvin Plantinga describes as a perfectly good, omnipotent, omniscient, immaterial creator (2000, 3). This is the way I use "God" in the present investigation. By contrast, "gods" are described by Clark and Barrett as including any supernatural intentional agent with whom we could interact, including such entities as ghosts, angels, demons, and ancestors (Clark and Barrett, 2011, 652). We shall see how Plantinga's faculty is depicted as a system which is capable of spontaneously and without inference providing beliefs about *God.* RE rules that such beliefs, when certain conditions are met, have warrant. By contrast, Clark and Barrett's faculty produces beliefs about a wide range of *gods*. Given that we have the faculty they describe, additional activities must be carried out to obtain beliefs about God. For example, subjects need to have the right cultural input, and to get God's properties right, they also need to reflect. This thesis relates to our discussion of belief-formation in chapter 3, where we saw that the cognitive processes CSR reveals contribute together with additional factors to produce beliefs about God, and that theologically correct beliefs about God are hard to form. Does CSR, by empirically supporting this view of belief-formation, empirically support central contentions in RE? I shall argue that it does not. The central contention in RE is that inferential and reflective activities are not necessary to form correct beliefs about God and thus (if God exists) obtain knowledge of God. But if, in the light

of CSR, this turns out to be necessary, then that means such beliefs do not have immediate warrant the way RE describes. Bringing CSR and RE together therefore leads to trouble.

We can think of this trouble in terms of a dilemma those who accept central contentions in RE and who find CSR-research compelling face. One can on the one hand base the interpretation of the god-faculty on Plantinga's thesis. I shall argue in section 2 that we can view this as a functionally specialized faculty for producing beliefs about God (i.e. the God of Abrahamic theism) rather than something else (such as spirits, ghosts, demons). The god-faculty in this sense seems capable of delivering the goods RE promises, namely non-inferential, properly basic theistic beliefs. However, a challenge arises given this interpretation. Section 3 argues that research in CSR fails to provide empirical support for such a faculty. One can on the other hand base the interpretation of the god-faculty on the thesis Clark and Barrett defend. I argue in section 4 that this is an empirically more plausible thesis. However, another challenge arises given this interpretation: the god-faculty, as mentioned, seems incapable of producing theistic beliefs the way RE describes. It needs support from various auxiliary activities which make beliefs about God inferential and therefore not basic the way immediate warrant requires. The upshot is that we cannot know God (if God exists) merely in virtue of the functioning of the god-faculty. To accept this means a departure from a central contention in RE concerning the immediate nature of knowledge of God. This is unacceptable for those who accept RE. In light of this, it seems that subjects must choose to opt either for RE or for CSR, since they cannot retain both. Given how CSR-research contributes to showing this, our conclusion is that such research undermines RE. We can relate this finding to our overall discussion of normative implications thus: since RE is a thesis or argument connected to theism, we here have an example of normative implications in sense (2) for what I call theist religion.

2. Alvin Plantinga and the god-faculty as a specialized system

> "God has so created us that we have a tendency or disposition to see his hand in the world about us. More precisely, there is in us a disposition to believe propositions of the sort *this flower was created by God* or *this vast and intricate universe was created by God* when we contemplate

> the flower or behold the starry heavens or think about the vast reaches of the universe." (Plantinga, 1983, 80).

Plantinga here describes circumstances and thoughts that trigger spontaneous formation of theistic beliefs. These beliefs can have *warrant* without independent support from evidence. This means a person can believe in a warranted fashion that God created the universe without having first been convinced by arguments. Warranted beliefs about God do not require such evidence. They require something much grander: God must exist. As mentioned above, Plantinga argues that *if God exists,* then he probably created us with a cognitive faculty that allows us spontaneously to form theistic beliefs in a way which suffices for warrant. This is the god-faculty or "Sensus Divinitatis" (Plantinga, 2000, 188-190).

Before we look closer at this faculty, let us first appreciate the central ideas of Plantinga's epistemology, known as *proper functionalism*. We return to this theory in the next chapter, where I argue in favour of a closely related theory. According to Plantinga's epistemology, a *properly basic* belief arises spontaneously and without inference from cognitive faculties that function properly. Such beliefs have warrant when the subject accepts them without basing them on anything. Most of our beliefs arise in this way, including perceptual beliefs, memory beliefs, a priori beliefs, and theistic beliefs (Plantinga, 2000, 178-179). The warrant of beliefs firstly requires that the responsible faculty functions or operates the way it *should*, given its function. This means it must not malfunction. Second, the faculty must operate in circumstances that are appropriate for it given its design. Beliefs that arise from the proper functioning of a faculty that operates outside of its appropriate domain of circumstances lack warrant (Plantinga, 1993, 4-7). Plantinga next suggests that the faculty must have a "design-plan" which specifies how it is supposed to operate. The design-plan for a faculty specifies what its function is. In short this means it specifies what it is for and how it is supposed to operate. The design plans for cognitive faculties specify that these systems respond to and process information and generate beliefs as output. Importantly, it is necessary for warrant that the belief arises from a faculty that operates in accordance with a design plan that aims at the production of true beliefs (rather than something else, such as to comfort us by making us overly optimistic). This means that a faculty whose design plan specifies that it aims at something else than true beliefs, but nevertheless often produces true beliefs, does not produce warranted beliefs (Plantinga, 1993, 11-17). This brings us to the final requirement,

reliability. The operation of a faculty for which all of the above is true must be such that there is a high objective probability of true belief once it is set in action (Plantinga, 1993, 17-18).

Central to Plantinga's epistemology of religion is the conditional claims that if theism is false, then theist beliefs probably lack warrant, and that if theism is true, then theistic beliefs probably have warrant. The basis for the first claim is that if God does not exist, then we lack the god-faculty and it is quite hard to see how theistic beliefs then could have warrant. For example, faculties or processes that bring about theistic beliefs would not meet the requirement of reliability. Moreover, our minds would not contain a faculty with a design-plan that specifies that production of *true theistic belief* is its constitutive aim. In contrast, if God exists, then it is very likely that we have a cognitive faculty that meets requirements for warrant. God loves us and would have wanted us to know him. He therefore would make sure that we have the requisite cognitive faculty to achieve this (Plantinga, 2000, 186-189). Such knowledge requires the faculty to give rise to *true theistic beliefs*. The design-plan for the faculty therefore will specify production of true theistic belief as its constitutive aim. If the system works the way God intended, then it will also accomplish its goal; it will be reliable the way warrant requires and thus enable us to know God (Plantinga, 2000, 179). Obviously, the faculty can only reach this goal if God exists. Moreover, not any belief will do. "God" means the omnipotent, omnibenevolent, omniscient deity of Abrahamic theism (see section 1). Therefore, the system must, to aim *at true theistic* beliefs, aim at God rather than other deities. Furthermore, the system must, to be reliable, form beliefs about God which get things largely right about him. Then, a faculty that spontaneously produces beliefs that get properties like omnipotence, omnibenevolence, or omniscience wrong will also not do.

What is the god-faculty Plantinga describes and how does it function? Plantinga characterizes it as a cognitive "input-output device" or "mechanism" that responds to a wide range of experiences, thoughts, and feelings, and generates theistic beliefs as output. This happens spontaneously, involuntarily, and non-inferentially (Plantinga, 2000, 174-175).[89] This means the god-faculty does not take beliefs as input and produce new beliefs about God as output. It neither produces beliefs about God through deliberate reasoning from evidence or arguments nor through spontaneous inferential activities (Plantinga, 2000, 175). Plantinga for

[89] This notion of a god-faculty is based on a thesis the 16th century theologian Jean Calvin defends about natural knowledge of God. Calvin called this knowledge the "Sensus Divinitatis" or our "sense of divinity". This idea indicates, Plantinga suggests, an instinctual human tendency or disposition to form beliefs about God (Plantinga, 2000, 170-171). Plantinga cashes out Calvin's idea in terms of a specific "faculty" ("mechanism", "device") which produces immediate beliefs about God using neither argument nor inference. This interpretation of Calvin has been challenged. Michael Sudduth (2009) argues that Plantinga's conception of an entirely non-inferential natural knowledge of God produced by such a faculty does not correctly capture Calvin's thesis (Sudduth, 2009, 60).

example says: “The deliverances of the *sensus divinitatis* are not quick and *sotto voce* inferences from the circumstances that trigger its operation” (Plantinga, 2000, 175). He later adds: “It isn’t a matter of making a quick and dirty inference...” (Plantinga, 2000, 176). According to Michael Sudduth, we can interpret Plantinga as contending that the natural knowledge of God the god-faculty (or sensus divinitatis) produces is completely non-inferential (Sudduth, 2009, 60). In that case, how do beliefs arise from the god-faculty? According to Plantinga, immediate beliefs about God, which are not based on other beliefs but instead represent points where thought commences, are “occasioned” in various circumstances. The circumstances whose occurrence can trigger the system, such that beliefs about God arise in such a fashion, include perceptual experiences of impressive mountains, the power of ocean-waves during a storm, but also the fragility of a flower, or the poetic dance of a leaf in the wind (Plantinga, 2000, 172-175). Other examples include feelings of thankfulness, guilt, or fear. Thankfulness as the result of a fortunate event can trigger the spontaneous belief that one should thank God. An experience of danger can lead to beliefs about God’s presence. Guilt as the result of bad behaviour can lead to the belief that one must ask God for forgiveness (Plantinga, 1983, 80-81). That beliefs like these are “occasioned” means that in a given situation they simply arise spontaneously without being based on perceptual experiences or beliefs (Sudduth, 2009, 84). These beliefs are not general beliefs in the existence of God but more specific beliefs about God whose truth entail the truth of theism in general (Sudduth, 2009, 84-85).

As I shall now argue, the god-faculty Plantinga describes meets plausible constraints for being a cognitive faculty or system. I do this before investigating the evidence for its existence. The epistemologist Jack Lyons (2009) suggests that a cognitive system is a functionally coherent mechanism specialized for handling a distinct set of tasks. This means such a system exhibits *functional specialization*. Lyons further points out that a cognitive system is not the same as a mere capacity. For example, we might have the capacity to recognize faces without having a specialized system for performing this task. A more general system that also solves other problems could be responsible. To say that we have a capacity for x is not the same as saying that we have a specialized system that does x. But a specialized system for x gives us a capacity for x (Lyons, 2009, 88-90, see also Boyer, 1994, 18-19). Plantinga is not merely suggesting that humans have the capacity to form theistic beliefs. That would be a trivial claim. He describes a faculty specialized to carry out this task. Functional specialization is tied to solving a restricted range of problems using a restricted set of procedures. Specialized cognitive faculties or systems exhibit what is known as *domain-specificity*. This means they do not operate on any kind of informational input or produce beliefs about anything using any kind of

procedure. Rather, they are specialized to handle specific types of information using specific procedures and to produce specific types of output as result (Sperber, 1996, 136 and Hirschfeld and Gelman, 1994, 3). For example, a face-recognizer responds to information relevant to the recognition of faces and produces beliefs about faces and facial expressions. A cheater-detector responds to cues that indicate if another subject is likely to defect from cooperation and produces signals about this. The god-faculty seems domain-specific in this sense. It produces beliefs about God, i.e. the Abrahamic God, rather than other deities (Clark and Barrett, 2010, 176).

Since the god-faculty is an input/output device that spontaneously and automatically produces *non-inferential* beliefs, the beliefs it produces are basic (not based on other beliefs), and if the subjects accept them as such, then these beliefs also count as *properly* basic and hence have immediate warrant. This is also the case with many other beliefs our minds produce, such as when we perceive things or employ our memory (Plantinga, 2000, 177-178). Consider how Plantinga describes the formation of theistic beliefs from the god-faculty:

> "{I}n the typical case we don't consciously choose to have those beliefs. Instead, we find ourselves with them just as we find ourselves with perceptual and memory beliefs." (Plantinga, 2000, 172-173).

This illustrates how theistic belief-formation is something immediate and non-deliberate. We do not choose to form a belief about God but simply realize that it has suddenly arisen. Moreover, Plantinga describes something that happens in normal individuals as the result of normal cognitive development, without the need for specific cultural inputs. He says for example that the development of the god-faculty in the individual only requires normal maturation. Moreover, it is typically in place in early childhood (Plantinga, 2000, 173). In light of these statements, the god-faculty seems to exhibit what Robert McCauley (2013) calls "maturational naturalness". Maturationally natural systems produce thoughts and beliefs automatically, without the need for conscious thought or exposure to culturally specific ideas. Furthermore, maturationally natural systems develop as part of normal cognitive development typically very early in life, without the need for specific instructions, training or culturally specific inputs. The naïve ontologies we discussed in chapter 2, section 4, which consist of intuitions about physical entities, biology, teleology, and agency exhibit maturational naturalness. Moreover, the disposition to acquire language or to learn how to walk is also maturationally natural (McCauley, 2013, 31-82). Based on the descriptions Plantinga provides

concerning the function and purpose of the god-faculty and given McCauley's account of maturational naturalness, we can plausibly think of the Plantingian god-faculty as a purported specialized cognitive faculty or system which is maturationally natural. The question we now shall ask is whether research in CSR indicates that we in fact have this system.

First, a brief reminder of the goal we pursue in doing this. My argument in the present chapter is that an advocate for the god-faculty thesis who accepts central tenets of reformed epistemology and who finds CSR compelling faces a dilemma. This advocate must either choose a Plantingian interpretation of the god-faculty or an interpretation based on the thesis of Kelly James Clark and Justin Barrett. We shall now see what challenges arise if he opts for the Plantingian interpretation.

3. The god-faculty dilemma horn 1: Alvin Plantinga

Given the above interpretation of the god-faculty and given Plantinga's claim that if God exists, then we have it, we obtain the thesis that the truth of theism leads to an empirical prediction concerning the makeup of human cognitive architecture. The prediction is this: among the many distinct faculties that make up the mind, we have a faculty specialized to produce true beliefs about God.[90] How can test this prediction? Do we in fact have this faculty?[91] Jack Lyons is right that questions about what cognitive systems or faculties we have and how beliefs arise are empirical and in principle open for science to answer. To set the tone for our discussion below let us appreciate what he says about the god-faculty thesis:

> "To figure out whether we do have a God module, we can do far more than make a priori guesses conditional on God's existence or nonexistence (or invoke the authority of Calvin, who was hardly renowned as a cognitive scientist). We can empirically discover (at least in principle) where these beliefs come from" (Lyons, 2009, 160).

[90] We cannot deduce the existence of this god-faculty from theism. Rather, the probability that we have this faculty is very high given theism (Plantinga, 2000, 189).

[91] Plantinga's conditional claim, if God exists, then we have the god-faculty, seems to be a material conditional. Such a statement is true if the antecedent (i.e. theism) and the consequent (the god-faculty thesis) are both false. Evidence that the latter is false therefore does not demonstrate that the conditional statement is false. Because Plantinga (2000) does not explicitly defend the truth of theism, it is not entirely clear whether he claims that we have the god-faculty. However, we will not go further into this interpretational issue, which is not important to the argument I present. The main task for us is to reveal what challenges that occur when bringing CSR and central claims in RE together.

What findings can we plausibly consider evidence? The above-specified requirements for being a cognitive faculty or system and Plantinga's depiction of the purpose and functioning of the god-faculty should guide us. This means we are looking for a functionally specialized, domain-specific system for producing theistic beliefs. Evidence for lack of functionality or lack of a domain for theism-related issues then violates the prediction. Furthermore, we are looking for a system which is automatic in functioning and arises as the result of normal maturation rather than deliberate training or exercise. In short, we seek a maturationally natural system. Then, evidence that there is a need for reflective thought, deliberate instruction, or cultural input to produce theistic beliefs violates the prediction. Finally, the system we seek aims constitutively at God and is reliable if God exists. Two factors bear on reliability here, (a) whether we form beliefs about the right deity (i.e. God), and (b) whether we get crucial properties of this deity largely right. Then, evidence for a susceptibility to form beliefs in other deities than God or in other supernatural agents and evidence for tendencies to get crucial properties of God wrong violate the prediction.

We start with functional specialization. On the current god-faculty thesis, the actions of God had specific consequences for the makeup of the human mind, the central one being that there is specialization in our minds for forming non-inferential and immediate beliefs about God. In virtue of being committed to methodological naturalism, empirical research on cognition does not bear directly on claims about the actions of God. But that does not stop such research from having strong relevance for their consequences, such as the one we currently consider. What does the evidence in CSR tell us about possible specialization for theistic belief-production? I have argued in favour of the by-product view of religion, according to which religious beliefs are by-products of the functioning of a variety of systems that evolved to solve tasks unrelated to religion. For example, we saw (Ch. 2, section 2) that Theory of Mind (TOM) evolved as solutions to challenges related to social interaction among conspecifics during the Pleistocene, and that numerous scholars in CSR claim that this system nevertheless is involved in religious belief-formation and behaviour presently. HADD evolved in response to challenges produced by the presence of predators and prey, yet this system can contribute to religious belief-formation (see Ch. 2, section 3, Ch. 3, section 5). The adaptive function of the attachment-system is to facilitate proximity between caregivers and offspring but when theistic concepts latch on to its functioning, it can help bring about theistic beliefs (Ch. 2, section 7). Similar stories can be told concerning other systems and processes. The main point is that religious beliefs do not arise from a system dedicated to produce such beliefs but as by-products

of numerous other systems when the subjects are exposed to cognitively attractive religious ideas. In light of this, we lack a specific cognitive domain for religion, not to mention beliefs about God (a specific deity). Some central scholars in CSR are quite adamant about this:

> "I have explained religion in terms of systems that are in all human minds and do all sorts of precious and interesting work, but were not really designed to produce religious concepts or behaviours. There is no religious instinct, no *specific* inclination in the mind, no particular disposition for these concepts, no special religion centre in the brain, and religious people are not different from non-religious ones in essential cognitive functions. Even faith and belief seem to be simple by-products of the way concepts and inferences are doing their work for religion in much the same way as for other domains." (Boyer, 2002, 378).

> "The mind has no department of religion, that is, the mind has no systems, structures, or processes specifically dedicated to managing religious materials" (McCauley, 2017b, 24).

The by-product thesis thus bears negatively on the thesis that we have the god-faculty by violating the functional specialization requirement. Notice that the clash is not (or at least not directly) with the thesis that *God* designed our minds with a functionally specialized system. The clash is with a consequence of the truth of that thesis, namely that our minds exhibit functional specialization for forming beliefs about God.[92]

We have seen (Ch. 2, section 2) that there are also scholars in CSR who favour adaptationist views of religion. These scholars might object to the claim that there is no functional specialization for religious issues. Johnson (2016) for example argues that religion is an adaptation whose adaptive function is to help us curb our selfish desires by way of making us afraid of supernatural punishment (see also the other examples of adaptationism in Ch 2). If this were correct, then there would in fact be functional specialization for religion. However,

[92] Notice also that the clash is not (not directly at least) with a general theistic thesis according to which God (somehow) uses human cognition to achieve various results (such as to give rise to belief in God). Even though it creates explanatory burdens, it seems possible to say of scientific findings that they (somehow) reveal the nuts and bolts of a divine plan. I am currently illustrating a clash between a specific consequence of the truth of Plantinga's god-faculty thesis and what current research in CSR indicates about the cognition responsible for religious thought. Are there other ways to resolve this conflict than by retreating from a specific god-faculty claim to a general accommodation-thesis? Here is a possible strategy. L.A. Hirschfeld and S. A. Gelman (1994, 8) suggest that in principle, a faculty (i.e. what they call a "module") could have subcomponents that themselves are faculties. To resolve the conflict, one could perhaps argue that CSR-scholars are only describing the substrate-faculties, and that we can interpret Plantinga's thesis as being about a sort of super-faculty which builds on those substrates while also exhibiting profound specialization in its own right. But then, the truth of Plantinga's thesis would still lead to empirical consequences. We could for example expect a distinct cognitive domain for theistic belief. Judging from the statements Boyer and McCauley make, there is no trace of those consequences.

that would not provide a plausible empirical basis for the specific god-faculty we are investigating. For example, the supernatural punishment system works by producing fear of punishment for selfishness and this involves beliefs about a variety of supernatural agents, not necessarily God. Moreover, whilst the god-faculty has a design-plan which specifies that aiming at God rather than other supernatural agents is the goal of this system, the plan for the supernatural punishment system, if it has any, seems to be to increase fitness by making us afraid of punishment for selfishness. That does not mean the system Johnson describes could not contribute to the formation of beliefs about God in the right cultural contexts. But the god-faculty is not like this. It is not a system evolved for other purposes (relating to fitness), which also gives rise to religious beliefs. It is specialized for doing the latter. Moreover, it does not merely contribute to form theistic beliefs. It spontaneously produces such beliefs. This example of adaptationism therefore does not empirically support the god-faculty.

We next turn to maturational naturalness. Face-recognition, walking, chewing, attributing mental states to other agents, or predicting the path of a moving physical object occur automatically and without conscious deliberation in all normal individuals. These activities and processes do not require prior instruction, specific artefacts, or exposure to culturally propagated ideas. This is because such activities and processes are maturationally natural (McCauley, 2013, 20-26). We have seen that the god-faculty also is natural in this sense. It is meant to develop as the result of normal maturation rather than specific instructions and activities, and it spontaneously produces beliefs about God in a wide range of circumstances. But in light of the account of belief-formation I have defended (Ch. 3), based on CSR, that is clearly not how theistic beliefs arise. As I argued in chapter 3 section 3, theistic beliefs arise in subjects who have a *gradually formed* and *culturally specific* belief-forming propensity. In such subjects, intuitive theistic beliefs arise from processes that involve cheater-detection, agency-detection, attachment, and so on, when the culturally specific concept of God first latches on to them.

As we have seen (Ch. 3), the importance of prior cultural exposure to explain religious belief-formation is something for example Justin Barrett and Jonathan Lanman (2008) emphasize. They argue that evolved systems and the underlying intuitions they produce can contribute to producing beliefs about God (and other deities) when subjects have the right "cultural scaffolding" and engage in belief-supportive religious practices. This is strongly at odds with the picture Plantinga depicts, where subjects spontaneously and without help from auxiliary belief-supportive activities or cultural input form the belief that *God created the*

universe and where becoming a theist seems to happen spontaneously rather than gradually the way I have argued.

We next turn to inferentiality. The system we seek produces theistic beliefs non-inferentially. It is not a reasoning device whereby subjects carefully consider evidence and base beliefs about God on it. Furthermore, (as illustrated above) it does not even involve what Plantinga refers to as "quick and dirty" inferences. Specific beliefs about God are simply "occasioned" in a variety of circumstances. Here again a conflict with CSR looms. Recall the three types of beliefs I specified in chapter 3, section 5. They all involve inferences. We saw first that intuitive theistic beliefs arise when the concept of God latches on to intuitive, inferential processes whereby the mind handles various kinds of information. We then looked at explanatory beliefs, which we interpreted as attributions, attempts to explain events. Recall for example the subject who formed beliefs about a storm being steered by some agency to destroy a house and later formed the explanatory belief that God was the responsible agent. I further suggested that this sort of belief readily gives rise to further beliefs, such as the belief that God in this way punished the inhabitants of that house. Such explanatory beliefs are not basic but based on prior beliefs and intuitions. Finally, folk-theological beliefs are based on religious teachings and therefore also inferential. Religious belief-formation is therefore inferential and for this reason different from how the god-faculty works. This further undermines the thesis.

We finally consider reliability. As mentioned above, we are looking for a system which is reliable if God exists. This means that (a) it gets things right by forming beliefs about the right deity, and (b) that it gets things largely right about this deity. Let us first discuss the latter.

A crucial property God has is omnipotence, meaning the ability to carry out all logically possible actions. The evidence in CSR indicates that when we think intuitively and non-deliberately about God, we get this property wrong. We are, as J. Barrett suggests, prone to attribute *limited* abilities to God (Barrett, 2002, 94-95). A study he did with Frank Keil provides evidence for this. Keil and Barrett tested how well subjects recalled stories involving God. They found revealing distortions. For example, subjects tended to (falsely) remember a story as saying that God had to finish listening to one prayer before he was able to attend another (Barrett and Keil, 1996). This indicates the presence of the intuitive thought that God is not omnipotent but has limited abilities. Another study, which also J. Barrett carried out, provides further evidence. Barrett here surmised that because we use cognition that enables us to represent humans when we represent deities, we can expect a bias towards "petitioning" (asking) God to do things we can do rather than things *only* God can do. This relates to the thesis (see chapter

2, section 3) that we tend to anthropomorphize God. To test this, J. Barrett distinguished prayers that request physical actions (putting out fires or repairing damaged things), biological actions (curing a sick person), or psychological actions (influencing someone's mental states). For God these things are equally easy to perform. In contrast, we can only do the latter type of action at a distance. J. Barrett proposes, based on the thesis that we anthropomorphize God, that there should be a bias towards *psychological* petitionary prayers. This turned out to be the case. Subjects were inclined to ask God to do things humans can do rather than things only God can do, as if not thinking of God as able to do literally everything (Barrett, 2002, 94-97). These two studies (together with the work of Stewart Guthrie, discussed in chapter 2) provide evidence for the thesis that we tend to anthropomorphize God, thereby getting one of his properties wrong.

A faculty which is reliable with respect to God must not only get his properties right. Such a profound system should also avoid mistaking other deities or ghosts, spirits, and so on, for God. For example, if God exists but we are prone to believe in Zeus or in ghosts, then (assuming the two latter do not exist), the responsible faculty is not reliable. It fails to get things right with respect to a somewhat broader range of related propositions about the supernatural. Our discussion in chapter 2, section 5 about our general susceptibility to MCI concepts is highly relevant to show that this is true of our minds. According to MCI theory, we are susceptible to MCI agents in general because of the cognitively attractive structure of such concepts. Among the many MCI concepts, we can include concepts of ghosts, spirits, elves, ancestors, angels, pixies, goblins, witches, trolls, and demons. Surely, the theist will also agree that many of these things do not exist. Furthermore, in light of what we saw in chapter 2 (section 6) on context-biases, we have grounds to believe that which specific entity subjects tend to form beliefs about depends strongly on the cultural context they are in. Assuming that many of the aforementioned entities do not exist, the responsible systems or processes do not seem reliable the way warrant requires. Their indiscriminate production of belief about such things indicates this. A profound specialized system to enable us to know God would do better. This further indicates that our minds are not reliable with respect to theism and the supernatural the way we could expect given the presence of the god-faculty.

Based on what we have seen, the current evidence in CSR strongly indicates that we lack a functionally specialized, maturationally natural, non-inferential god-faculty. Our imagined god-faculty advocate, who is committed to RE and interested in accommodating the evidence in CSR therefore encounters trouble by opting for the Plantingian interpretation of the god-faculty. Before looking at the other main interpretation, let us first consider two responses.

It has been argued that psychological evidence suggests that children spontaneously attribute omniscience to God, thereby getting at least this crucial property right. Let us now discuss whether this indicates that human minds are more reliable with respect to God's properties than what I have currently argued. According to Justin Barrett, attribution of omniscience or "super-knowledge" to God happens spontaneously and without much effort, even in children (Barrett, 2004, 75-80). J. Barrett and Rebekah Richard (2003) suggest that at a very early age, children develop a flexible notion of agency that applies to human and non-human agents and even to God. The cognitive default-assumption children make when employing this notion is that agents have *infallible* beliefs. J. Barrett and R. Richert refer to their view of these matters as "preparedness" (2003, 301). Presumably, the idea is that children's default-cognition prepares them *for becoming theists* by nudging them to attribute epistemic infallibility to God. In support of their view, Barrett and Richert point to a study that involves a false-belief task modified to deal with theistic beliefs. In this study, Barrett and colleagues presented children with a cracker-box containing rocks and then asked the children what their mother and God believe the box contains. The issue is whether the children would attribute a false belief by answering that their mothers and God believe there were crackers in the box. Barrett found that at the age of 3-4, the children attributed the belief that the box contained rocks to both their mothers and to God. Children above 5 attributed false beliefs to their mothers by answering "crackers" but failed to attribute false belief to God. It seems they thought that God would know there were rocks in the box even though they were able to attribute false beliefs to him. Barrett and Richert interpret this finding as supporting preparedness. On their interpretation, the older children realized that God is infallible yet attributed fallibility to their parents. Furthermore, very young children got things right with respect to God by assuming that God (and every other agent) is infallible. Theological accuracy was thus achieved because God *is* infallible (Barrett and Richert, 2003, 305-308).

Did the children intuitively and without deliberation actually realize that God knows *everything*? The first reason to be skeptical of this is that, as De Cruz and De Smedt (2015, 49) point out, failure to attribute false beliefs to an agent is not the same thing as attributing omniscience[93] to that agent. We can interpret the latter as perfect knowledge of all facts and states of affairs. In contrast, what J. Barrett and Richert call infallibility seems to be something like perfect reliability, i.e. that one only forms true beliefs. One might be epistemically infallible with respect to a limited or a broad range of issues. An agent infallible with respect to a limited

[93] Presumably, "super-knowledge", which is the term Barrett uses, means something close to omniscience.

range of issues is perfectly reliable with respect to those issues but not omniscient. Only infallibility with respect to *all* propositions and states of affairs entails omniscience. It is not clear, based on the study discussed above, that this is how the younger or older children thought of God. Therefore, even if the evidence indicates that the children thought that God is epistemically infallible, that does not show that they thought of him as omniscient.

Second, another interpretation of the results, which is consistent with the anthropomorphism thesis, is available. There is evidence that subjects represent God as an agent with access to what Boyer (2002) calls "strategic information" rather than any kind of information (Purzycki et al. 2012). Recall (chapter 2, section 6) that such an agent has full access to socially relevant information, that is, information that triggers cognitive faculties that govern social exchange. Information about moral transgressions is strategic and this includes information about attempts to *deceive* others.[94] In the study above, the children were asked to imagine that their mom would enter the room and encounter the cracker-box (with rocks). Would the outward appearance of the box *fool* her to believe that it contained crackers? Arguably, being asked to consider this could have led the children to think that presenting this box to mom (or other agents) was an attempt to deceive. This would make the information about the content of the box strategic. God, an agent with access to such information, would know what the box contains. On this interpretation, the children did not fail to attribute a false belief to God because they intuitively thought that God is omniscient. Rather, they thought of God as having access to a particular piece of strategic information. It is important to stress just how different an agent with access only to such information is from an agent who literally knows everything. Consider the set that contains all logically possible facts. Then compare this with facts about human social existence. An agent with full access only to strategic information has access to all the latter facts but nothing else. However, the set that contains the latter facts is surely an unimaginably small subset of the former set. Then, given that the children only attributed knowledge of the contents of the subset but not the former set, they did not think of God as omniscient. This interpretation of the results allows us to account for the results without denying the thesis that we intuitively tend to anthropomorphize God.

---

[94] In a study of how the epistemic states of deities are mentally represented, Purzycki et al. measured how long subjects took to respond to questions about what deities know about strategic and non-strategic information. They found that subjects processed statements about the former faster than the latter. For example, on average, they answered questions about what God knows about human moral behavior faster than questions about what God knows about morally irrelevant facts. Furthermore, Purzycki et al. found that subjects were especially fast when processing statements about negative moral behavior, such as statements about moral transgressions (Purzycki et al. 2012, 850-855).

Let us finally consider an entirely different response to my claim that CSR strongly undermines the thesis that we have the god-faculty. To appreciate this response, we need to know something more about the thesis Plantinga defends. Plantinga distinguishes what we can expect if mere theism is true and what we can expect if a more complex Christian version of theism is true. Assuming the former leads to the expectation that we should have a god-faculty that enables immediate and spontaneous knowledge of God. We just saw that CSR violates that expectation. But given that Christianity is true, what we can expect changes. According to Christianity (the way Plantinga presents it), *original sin* has infected us, and this has negative cognitive consequences. Sin makes us proud and selfish, there is a tendency to form false beliefs about what God is like (for example that he is indifferent or bad), and about what we are like (for example that we are better than other people and that our moral failings are not that bad). Sin thus affects our belief-formation. Nevertheless, a sin-damaged faculty enables a rudimentary grasp of God (Plantinga, 2000, 199-216). One might argue that these sin-induced effects are consistent with the evidence discussed above and that by introducing sin as an auxiliary hypothesis one can therefore salvage the thesis. Does that work?

Notice that the thesis that we *do not have* a specialized system devoted to producing beliefs about God and the thesis that we have a functionally specialized, albeit damaged god-faculty are still at odds. But, admittedly, the empirical consequences of the truth of the latter thesis might be easier to square with the above-reviewed evidence. It seems at least possible that tendencies to form incorrect religious beliefs could be the result of sin. But does introducing sin help the god-faculty advocate? As for example John Teehan points out, an evolutionary study of the origins of humans seriously challenges this doctrine, which supposedly involves claims about how our present state was preceded by a morally superior state from which we somehow "fell" (Teehan, 2016a, 50, see also Teehan, 2018 for more extensive discussion). Then, if sin is to be of assistance, one must propose an interpretation which is not undermined by lack of evidence for a "fall" in human natural history. I am not denying that this can be done. The question is whether introducing sin as an auxiliary hypothesis to handle the evidence above is of any help. In relation to the "cognitive problem of evil" (see Ch. 4) I did not argue that the findings of CSR concerning our susceptibility to evil necessarily rebut theism or make accommodation-theses impossible to sustain. The main point was that these findings create a need for additional explanation. That also seems true of the findings we reviewed above. The Christian notion of original sin can be put to service to help the advocate handle the empirical evidence discussed above, but this service is not without cost. By introducing sin, one brings forth a barrage of further Christian claims which also needs to be defended. Sin after all is a

central Christian notion. It is arguably at the heart of much Christian theology (and strongly related to the doctrines of the atonement and resurrection of Jesus). But then, notice that the god-faculty thesis begins to be quite complicated. Our god-faculty advocate, who accepts Plantinga's thesis, believes we have the god-faculty, but feels troubled by the seeming counter-evidence in CSR therefore does not get an easier job by introducing sin as an auxiliary hypothesis.

In light of this, the conclusion stands: the evidence in CSR is strongly at odds with what we can expect given the truth of the thesis that we have a Plantingian god-faculty. The advocate for the god-faculty thesis therefore should consider pursuing another strategy. Given that he finds CSR-research convincing yet also will not budge in his commitment to central tenets of reformed epistemology, this seems necessary. Is another plausible construal of the god-faculty at hand, which is in line with the best evidence but at the same time retains important RE-tenets? We now turn to a proposal that purports to deliver the goods our advocate needs.

4. The god-faculty dilemma horn 2: Kelly James Clark and Justin Barrett

We saw in section 1 that Kelly James Clark and Justin Barrett contend that CSR and RE both indicate that religion is cognitively natural but not in exactly the same way. While the former sees beliefs about "gods" as natural, the latter contends that it is the belief in God (presumably the Abrahamic deity) which is natural. I have argued that the latter view faces a challenge. CSR does not support the claim that belief in something as specific as God is natural (in the sense of arising spontaneously and without inference from a cognitive faculty). But must we interpret the god-faculty as a specialized system God rigged to produce beliefs about himself and nothing else? Clark and Barrett point to another possibility, inspired by Jean Calvin. Calvin in their view held that God has endowed us with a general sense of his divinity. This natural knowledge of God can then be expanded when the subjects appreciate how God is manifested in creation (Clark and Barrett, 2010, 175, cf. Sudduth, 2009, 62-63). Clark and Barrett argue further that the general sense of divinity Calvin described might come to expression in the form of beliefs in "gods", that is, a variety of supernatural agents, such as deities, spirits, ghosts, angels, and so on (2011, 649). Here CSR enters the picture: This thesis about religious belief is supported by CSR, which therefore points us to a god-faculty-thesis:

> "There now seems to be good empirical reason, provided by cognitive scientists studying religious thought, to believe what some philosophers and theologians affirmed on theological grounds: that we have a maturationally natural god-faculty, although "religious faculty" or sensus divinitatis may be more precise and relevant terms" (Clark and Barrett, 2011, 649).

Notice how strongly this seems to differ from my contention that CSR *undermines* the thesis that we have a god-faculty. We need to appreciate, however, that the term "god-faculty" means very different things in the above and in the current thesis:

> "By "god-faculty", we mean that the ordinary arrangement and functioning of cognitive architecture in human minds often produces nonreflective, unreasoned belief in gods. By "gods", we refer to any supernatural intentional agents whose existence would impinge upon human activity. We are not arguing that this god-faculty is a dedicated functional system, a special add-on to human minds, nor that it is divinely implanted by natural or other means." (Clark and Barrett, 2011, 652).

In a discussion of what sort of faculty this is, Clark and Barrett (2011, 649-654) describe a range of underlying faculties and cognitive processes of the kind we have discussed in chapter 2 (Theory of Mind, HADD, teleology, and so on) and argue that their interplay provides the basis for this god-faculty. This faculty is thus an emerged by-product of our cognition (Clark and Barrett, 2011, 652). How does this system work and what can it achieve in terms of belief-production? As is clear from the above statements, the god-faculty operates in a non-reasoned and automatic way, providing beliefs without the use of reflection. It seems to operate in a non-inferential and immediate manner (see Clark and Barrett, 2010, 174). Clark and Barrett describe the god-faculty as achieving something profound in virtue of spontaneously providing immediate and non-reasoned beliefs about "gods". To form these spiritual hunches or "inklings" represents the beginning of a larger spiritual journey: These sorts of religious hunches help bring about the realization that a divinity exists (Clark and Barrett, 2010, 187). This sense of divinity is very general and unspecific and so needs to be refined and elaborated, something which might require reflective activity (Clark and Barrett, 2011, 666). Presumably, what Clark and Barrett mean by this is that by *reflection* we can close the gap between the coarse hunches the god-faculty produces and the belief in God, which is where we ultimately are heading. As we saw earlier, RE attempts to show how knowledge of *God* (rather than other supernatural agents) is possible. But this goal will be achieved in different ways through the god-faculty we discussed previously and through the current faculty. With Plantinga's god-

faculty, the route to the belief in God can be traversed spontaneously and without inferential activity. By contrast, with Clark and Barrett's faculty, less can be achieved in terms of spontaneous cognition, and a more complicated route via beliefs in "gods", the realization that some kind of divinity exists, and subsequent reflection must therefore be followed. Let us now investigate whether this proposal represents a better alternative for our god-faculty advocate, who by stipulation commits to both RE and CSR.

We start by considering the empirical evidence. In light of our discussion of the evolutionary origins of religion (Ch. 2, section 2), it is plausible that our susceptibility to belief in "gods" is a by-product of how many parts of our mind work rather than being the product of one specialized system. Moreover, I argued (Ch. 2, section 5) that we are susceptible to belief in MCI beings in general because of the cognitive attractiveness of such notions. MCI beings, that is, intentional agents with properties that minimally violate naïve ontologies, seem related to what Clark and Barrett mean by "gods". The cognitive attractiveness of these notions plays an important part in the cultural epidemiology Dan Sperber and Pascal Boyer (and others) support (Ch. 2, sections 4, 5). Such notions are culturally ubiquitous because of how cognitively attractive they are. Clark and Barrett use the term "god-faculty" as a name for the phenomenon that our minds generate and easily transmit such ideas, thus giving rise to such cultural products. It seems also that the god-faculty, or in other words, the above-mentioned susceptibilities, can help explain (proximately) how specific MCI-beliefs arise in a given context, if we assume that subjects are exposed to the right MCI concepts in that context. Recall (Ch. 3, section 5) for example J. Barrett's discussion of "Doug", the subject who remarkably escaped an accident in a silo. It could be that the underlying HADD-component of the god-faculty gave rise to intuitions about agency during the event. Doug later obtained a religious concept by talking to his friends after the accident and then used that concept when explaining the event. Here then, the god-faculty helped bring about a specific religious belief in a given context where there was exposure to a religious idea. The god-faculty Clark and Barrett describe seems to require such exposure in order to give rise to specific MCI beliefs. Furthermore, we saw above that it allegedly can contribute to belief in the divine and ultimately also facilitate the realization that God exists. One might think of this in ultimate terms as a thesis about a gradual spiritual maturation in humanity whereby the god-faculty brings us some distance but whereby our reflective thought and the development of religious belief-systems and practices must step in to bring us all the way. But we can also, it seems, think of this in proximate terms as the thesis that the god-faculty to some extent helps bring about belief in God. It does not, like Plantinga's faculty, trigger this belief directly. Instead, by triggering unreasoned thoughts about "gods" in

many cases, the god-faculty nudges subjects to commence on a personal path which gradually leads to that belief.

Clark and Barrett argue that beliefs about gods (MCI beings) arise non-inferentially in an automatic and immediate sense (2011, 174). Presumably, by this they mean that no conscious or deliberate inference is involved. As illustrated earlier in the chapter (and discussed in Ch. 3), non-conscious inferential activities are central to the production of beliefs in the supernatural. Justin Barrett and Jonathan Lanman also strongly emphasize this. Prior non-reflective, inferential activity can make general beliefs in the supernatural plausible (Barrett and Lanman, 2008). In light of this, theistic beliefs and other religious beliefs that arise from the god-faculty seem to be inferential in the sense that non-conscious inference is involved. But reflection, an example of conscious and explicit inference, is not involved. That sort of activity therefore seems needed to move from the beliefs in gods to, presumably, the goal of a obtaining a correct understanding of God. In light of this, the current god-faculty thesis is empirically plausible. It uses faculty-terminology to discuss non-conscious processes strongly related to those we have discussed in the present investigation and appeals to empirical evidence in CSR to support it. Since CSR supports the claim that we do have the belief-forming susceptibilities Clark and Barrett describe under the heading of "god-faculty", our imagined advocate here seems to avoid the empirical problem which arose in relation to the Plantingian thesis. Is then Clark and Barrett's god-faculty the solution to his trouble? To find out, we must investigate whether one can accept it whilst at the same time retaining central tenets in RE.

A serious challenge here arises. RE strongly emphasizes that we can know God *without* the use of inference or argument. RE offers a particular thesis about cognitively natural knowledge of God, which is related to the idea that beliefs about God can be properly basic. As we saw in section 2, RE excludes both non-conscious ("quick and dirty") and conscious (reflective) inference from proper basicality. As Michael Sudduth points out, RE here offers a particular interpretation of reformed thought, according to which knowledge of God can be achieved non-inferentially and without conscious reflection (Sudduth, 2009, 77-95). A serious problem looms here concerning how to close the gap between belief in what Clark and Barrett refer to as "gods" and the belief in God without violating this specific thesis in RE. Arguing that the god-faculty merely helps nudge subjects towards the belief in God does not solve the problem because if this faculty must have help from various inferential and reflective activities to yield that belief, then we cannot know God in the immediate way RE contends. Given our discussion of the present god-faculty, it clearly seems insufficient to yield such a result. Both non-conscious inference (carried out by evolved systems) and explicit reflection seem

necessary to close the gap between "gods" and God. This seems to be what Clark and Barrett (2011, 666) indicate when pointing to the need for reflection. It then follows that belief in God is not properly basic and that we cannot know God without inference or argument. Our god-faculty advocate therefore encounters a serious problem by opting for Clark and Barrett's thesis. By departing from Plantinga to accommodate CSR, he ends up departing from a central tenet of RE. For this reason, despite how empirically plausible the thesis is, it fails to solve his specific problem.

Let us consider a response to this argument. Clark and Barrett might object that the gap between unreasoned religious gut-feelings and belief in God is not as large as I assume. Perhaps then, the gap between the output of the god-faculty and the belief in God can be closed in ways RE allows as compatible with immediate, properly basic knowledge. Clark and Barrett for example suggest that God might allow our cognitive mechanisms to produce beliefs in MCI agents to make us aware of the existence of the divine.[95] In virtue of forming beliefs about such agents, one gets a "taste of divinity", a sort of "core belief" in a morally good "super-knower" (Clark and Barrett, 2011, 667). Clark and Barrett also suggest that these beliefs, which even include beliefs in elves, ghosts, and fairies, might just represent a "surface impurity" in the general sense of the divine (Clark and Barrett, 2010, 187). Perhaps then a susceptibility to MCI beliefs already is belief, albeit in rudimentary version, in God? Then there is no gap to speak of which we only can traverse through the sort of reflection that would make theistic beliefs inferential and therefore non-basic.

However, this response fails. It is not plausible that a sense of God's divinity, even if only rudimentary, can consist in a susceptibility to form beliefs in MCI agents. Consider on the one hand elves, ghosts, or fairies, minor supernatural beings who might play in gardens during the night or haunt basements and abandoned buildings. Consider on the other hand God, the almighty creator of the universe. A susceptibility to believe in the former is not plausibly even the beginning of a sense of God's divine majesty. It seems simply to be a susceptibility to believe in something other than God. The following scenario therefore seems possible: I strongly believe that ghosts haunt my basement and that fairies play in my garden at night, but entirely unbeknownst to me a morally good super-knowing divinity exists. The MCI beliefs I form do not make me aware of the existence of that divinity and do not constitute a belief in this divinity. The gap, which must be traversed by reflection, therefore remains. Moreover, having these MCI beliefs does not seem to offer much spiritual help. From a theist point of

[95] But it clearly leads to further questions. Why did God not just make sure we obtained correct beliefs about him right from the start? Why allow all this confusion and quarrel about the supernatural to burden our existence?

view, they might even represent a superstition which lures me *away* from realizing that God exists. In conclusion, the god-faculty thesis Clark and Barrett defend does not solve the problem our imagined advocate faces. In this case, the problem is not lack of evidence in CSR but the need to depart from RE in order to accommodate that evidence.

In conclusion, the god-faculty advocate faces a dilemma from which it is hard to escape. We have seen that he has two primary interpretative options. By construing the god-faculty in a Plantingian fashion, the thesis leads to a clash with CSR. By construing the god-faculty the way Clark and Barrett do, the thesis leads to a departure from the central RE-claim that we can know God without inference or argument. Given that our imagined advocate finds CSR convincing and will not budge in his commitment to RE, neither of these options work. We can conclude that CSR bears negatively on the thesis that we have a god-faculty. Given that this is a central thesis in reformed epistemology, CSR undermines reformed epistemology. We then have another example of normative implications for what I call theist religion in sense (2): CSR undermines an argument or thesis connected to theism.

Conclusion

We have in the previous and in the present chapter seen how input from CSR bears in assessment of theistic arguments or theses, and that such research therefore has normative implications for what I call theist religion in sense (2). Philosophy of religion can fruitfully employ insights from current cognitive science and is therefore not solely an enterprise one should carry out from the armchair. There is a pattern here we should appreciate. It is perhaps not overly surprising. When theistic arguments or theses make claims or entail claims about human psychology, findings in cognitive science bear strongly on them. Given that philosophy of religion also contains many investigations that deal with purely abstract matters or empirical matters beyond the scope of cognitive science, CSR does not seem to have extensive (wide-ranging) relevance in this area in general. Nevertheless, the impact can be strong in particular cases. We found that John Teehan's argument reveals that it is hard in the light of the cognitive science of evil to maintain that God sculpts our cognition. We then found that CSR creates problems for at least two moral arguments for the existence of God. In the present chapter, I argued that such research is at odds with reformed epistemology. These findings indicate that CSR has an impact in philosophy of religion which is strongly negative in particular cases. In conclusion, based on what the previous and the present chapter has argued, CSR has only moderately extensive but strongly negative impact for theistic arguments and theses.

This concludes part 2 of the present investigation. The third part, which consists of chapters 6 and 7, explores normative implications in sense (3). I ask what relevance CSR has in evaluations of the epistemic status of *theistic beliefs*. I pursue this issue by arguing in favour of an externalist, virtue-oriented conception of justification. We will then see that CSR sheds light on central factors that bear on the epistemic status of beliefs in accordance with that theory and therefore has the implications we seek.

Chapter 6. A Virtue-epistemology for the Theistic Disposition

1. Introduction and chapter-outline

The previous chapter argued against the contention in reformed epistemology that we have a god-faculty. As far as we can tell in the light of CSR, we have no such faculty. Instead, the evidence indicates that theistic beliefs arise from a culturally contingent belief-forming disposition. What epistemic status does this disposition have and in what way can input from CSR contribute in an evaluation of this? The main task of the current and the next chapter is to answer this question. We thus shed light on how CSR has normative implications in sense (3).

The current literature (see chapter 1, section 5) contains three main positions on the implications CSR has for theistic beliefs: neutrality, positive relevance, and debunking. Debunking is often suggested as a possible worry or problem, but then it is argued that CSR has no or very limited epistemic relevance. This is the "epistemic neutrality thesis", ENT. I presented (Ch. 1, section 5) an internalist and an externalist argument in favour of ENT based on claims we find in the present literature. A central premise in the internalist argument (I-ENT) is the claim that only factors that are internal have epistemic relevance. The second premise states that CSR does not shed light on any such factor. I argued (chapter 4) that there are strong grounds for denying the central premise and accepting externalism. Furthermore, CSR bears on at least some factors that are epistemically relevant according to internalism, namely arguments for or against the existence of God. We thus also found strong reasons to deny the second premise. This showed that I-ENT fails. The externalist argument for ENT has the plausible premise that facts about belief-forming processes have epistemic relevance. The central premise states that CSR *cannot* explain how theistic beliefs arise. Chapter 3 countered this by arguing that such research *can* explain this. Both methodological issues about explanation (section 2) and the account of the theistic disposition and the formation of three types of theistic beliefs (sections 3, 4, and 5) indicate this. So, two central arguments for ENT fail, and CSR potentially has epistemic relevance for theistic beliefs. It remains to see what that relevance consists of. Is it positive or negative? We have just found (chapter 5) that the positive relevance view is

problematic, since it involves the problematic notion of a god-faculty. The remaining position to consider is debunking.

In recent discussions about the implications of CSR, we find many criticisms of this view. Consider for example the argument Justin Barrett (2007) presents (and criticises). It draws the conclusion that theistic beliefs are irrational on the basis of the by-product thesis. As another example, Aku Visala (2011) and Michael Murray (2010) criticize a debunking-argument involving HADD and its alleged unreliability. Critically scrutinizing such debunking-arguments is important. However, this criticism does not establish the falsity of the debunking position, which also can be defended in other ways. Consider for example the argument John Wilkins and Paul Griffiths (2012) propose. They argue that an evolutionary understanding of human cognition found in CSR-scholarship gives us grounds to believe that human cognitive faculties track fitness-relevant truth but fail to track *religious truth*. This is because we lack grounds to believe that successful tracking of religious truths has fitness-value. In short, there is no link between getting our deities right and increasing our fitness. This allegedly undermines religious beliefs (Wilkins and Griffiths, 2012, 133-146, cf. Joyce, 2006 for a similar argument concerning evaluative beliefs). But though this is suggestive, more needs to be said. Wilkins and Griffiths are for example silent about which theory in CSR they find correct and what the details of that theory could tell us about religious beliefs and their formation. Their argument concerns natural selection and the possible fitness-value of getting things right in a given domain (such as common-sense, morality, and religion). This means the focus is largely on *ultimate* questions (see Ch. 3, section 2). What an ultimate evolutionary account tells us about truth-tracking in a domain is epistemically relevant for beliefs in that domain, but *proximate* details that show us that the likely belief-forming process *fails to get things right* will better substantiate a negative conclusion. Furthermore, there is also a need for a defensible normative principle that rules against justification in cases where there is no truth-tracking. The next chapter presents two arguments that provide what is needed. I call the first *the argument against epistemic virtue*, and I call the second *the argument against epistemic justification.* The account of the theistic disposition in chapters 2 and 3 (and to be further developed in the next chapter) plays an important role in these. I shall show that we have grounds to deny the view that this disposition is *reliable* in the right way. The beliefs it gives rise to therefore lack (doxastic) justification. The present chapter prepares the ground for these arguments by defending a normative theory on the basis of which we can make that ruling. This is an externalist, virtue-oriented, epistemological theory known as *agent-reliabilism*.

Here is an outline of the chapter. Section 2 provides a brief introduction to virtue-epistemology (VE) and the notion of an epistemic virtue in general. I argue against Linda Zagzebski's (2008) version of VE, which builds the conception of epistemic virtue on Aristotelian moral virtues. By making the presence of epistemic virtue in a robust Aristotelian sense necessary for knowledge and justification, this theory is too restrictive. I point to agent-reliabilism as a more plausible version of VE. Section 3 presents agent-reliabilism in more detail. In this theory, a central contention is that epistemic virtues are examples of cognitive abilities, dispositions to achieve things in the domain of knowledge. Such abilities can be grounded in a broader understanding of ability in general. Furthermore, and important for our purpose of assessing theistic beliefs, on this theory the epistemic justification of beliefs hinges on whether they arise from a process grounded in an epistemic virtue. Section 4 argues that agent-reliabilism is preferable to two closely related theories, process-reliabilism and Plantinga's proper functionalism. These three theories share the important view that facts about belief-formation strongly affect the epistemic status of beliefs, but agent-reliablism has an advantage. It is superior to process-reliabilism in virtue of how it handles counterexamples according to which reliability is insufficient for justification. Proper functionalism can also handle such examples. Nevertheless, agent-reliabilism has an advantage compared with proper functionalism in denying Plantinga's requirement that our cognitive faculties must have a design-plan according to which truth-aiming is a constitutive goal. This is too restrictive: It rules out justification in cases where there is no such design-plan but where justification intuitively seems to obtain. In some cases, it is also too permissive. It allows us to attribute justification in cases where this is not plausible. Finally, section 5 clarifies how the theistic disposition can be viewed as a candidate epistemic virtue whose epistemic status has implications for the beliefs that arise from it. I point out how the theistic disposition is a gradually acquired habit which involves the operation of evolved faculties as well as their entanglement with the concept of God, a culturally specific, cognitively attractive idea in Abrahamic traditions. The upshot is that the theistic disposition is a hybrid candidate virtue, which combines what R. McCauley (2013) calls "maturational" and "practiced" naturalness. The ground is then ready for the investigation in the next chapter, which leads to a conclusion concerning the implications CSR has for the epistemic status of theistic beliefs.

2. Two types of virtue-epistemology

Virtue-epistemology (VE) differs from other forms of epistemology by focusing on stable features of the subject who forms belief instead of focusing only on the beliefs that the subject forms. VE thus resembles virtue-ethics, which focuses on the subject who acts and derivatively on the moral status of the acts the subject carries out (Zagzebski and Fairweather, 2001, 3-4). The central notion in VE is that of an intellectual or epistemic virtue. This notion relates to the philosophy of Aristotle. Aristotle as well as other philosophers of his age saw virtue (arête) as a kind of excellence that anything with a function can have. According to this kind of view, for an x to be virtuous is for x to perform well or function excellently qua the kind of substance it is. This applies to humans and functional parts of humans (such as eyes or ears), and it applies to artefacts like tools or furniture (Greco, 2007, 177). But Aristotle also had a more specific theory of virtues. Book 2 of Aristotle's *Nicomachean Ethics* provides good examples of an appeal to virtues in the domain of morality as well as that of rationality and knowledge-obtainment. Humans that live a good life have according to Aristotle realized *moral virtues* such as fairness, generosity, courage, and temperance. These character-traits arise gradually as the result of frequent repetition of what counts as a right action in a given circumstance for a given agent. Gradually, by prudently making the right choices in various situations over time, behavioural and attitudinal proclivities come about as stable features of the subject. Such proclivities make moral action natural in the sense that they involve a disposition to act in ways that are morally right for a given subject in a given circumstance. In book 6 of the *Nicomachean Ethics,* Aristotle also describes *intellectual virtues*. These are achievements of the subject in various domains of rationality. One intellectual virtue is practical rationality (prudence). This virtue is involved in the formation of moral virtues because it is involved in enabling the subject to rationally choose the right courses of action in specific circumstances. Another virtue is scientific knowledge.[96] This refers to demonstrative or explanatory knowledge of the most basic aspects of reality.

Virtue-epistemological theories tend to combine the element of a disposition to achieve things and the element of actual achievement. Achievement here means the attainment of true belief in a way that counts as epistemically good. Theories in this area differ, however, when it comes to how they envisage epistemic virtues. We can loosely distinguish *character-based* theories, which model the notion of epistemic virtue on Aristotelian moral virtue, and *faculty-*

[96] In addition, Aristotle also describes the ability to produce, i.e. technical knowledge as an intellectual virtue.

based theories, which see epistemic virtues as reliable belief-forming faculties or habits. VE started as an example of the latter when Ernest Sosa introduced the notion of an intellectual virtue into contemporary epistemology, but eventually also the other version of such epistemology emerged (Zagzebski and Fairweather, 2001, Sosa, 2017, 140).

Linda Zagzebski's epistemology is an example of character-based VE, strongly based on Aristotelian virtue-ethics. On her theory, knowledge is the attainment of true belief in a way that involves what she calls intellectual virtue. This makes the attainment of knowledge a moral matter. A virtue in general is an acquired excellence of the person, which involves a motivation to succeed and a tendency to actually succeed (Zagzebski, 2008, 443). Intellectual virtues are excellences to perform well in the domain of knowledge-attainment. They involve a motivation to obtain knowledge (rather than other things). Such excellences are, together with other moral virtues, *character-traits* that subjects acquire gradually as the result of effort and practice. Virtues thus become established as lasting traits through a sort of habituation-process. Examples of intellectual virtues include intellectual courage, open-mindedness, and epistemic responsibility. These traits not only involve a motivation to obtain knowledge but also a reliable disposition to obtain it. Virtue is, as Zagzebski puts, it a "success-term". Virtuous persons tend to achieve what they are motivated to achieve (Zagzebski, 2008, 442-444, Zagzebski and Fairweather, 2001).

Roberts and Wood (2014) view Zagzebski's character-based VE as providing a plausible account of what they describe as "high-end" forms of knowledge, such as those related to making scientific discoveries or obtaining subtle forms of moral self-knowledge. The account is in their view plausible given its emphasis of character-traits like intellectual courage and intellectual humility, and the importance of being motivated to obtain truth. These factors seem to provide sufficient conditions for high-end forms of knowledge (Roberts and Wood, 2014, 10-11). But by making such factors necessary, the theory encounters a serious problem, since many cases of what we plausibly can view as knowledge do not involve such factors. Roberts and Wood provide the following example. Consider a subject sitting in a well-lit room. Suddenly all the lights go out. In this case, one hardly, Roberts and Wood point out, needs to be intellectually courageous or open-minded to obtain knowledge that the lights went out (Roberts and Wood, 2014, 10-11, see also Sosa, 2017, 147-148). This seems correct. It is too restrictive to put forward the presence of intellectual virtue (in the above sense) as a necessary requirement for knowledge. Not all cases of knowledge plausibly involve virtue in that specific sense.

Ernest Sosa (2007) and John Greco (2010) argue in favour of the other main type of VE, a faculty-based version that focuses on the reliability of cognitive faculties or habits. This version of VE rejects the thesis that knowledge and justified belief necessarily require a motivation to obtain knowledge or the habituation of praiseworthy character-traits like those mentioned above. Sosa and Greco argue in favour of theories where the crucial element is achievement of the sort we best explain as a manifestation of an epistemic virtue. They argue that dispositions to form beliefs can be epistemic virtues as long as they meet certain requirements, the central one being reliability. Furthermore, such dispositions have their basis in the natural functioning of cognitive faculties or in gradually acquired thought-habits. This version of VE provides a plausible account of knowledge and epistemic justification. By not making the manifestation of moral character-traits necessary, this theory has an advantage over Zagzebski's theory in being more permissive, thus capturing the many cases of mundane knowledge.

What about sufficient conditions? Roberts and Wood grant that faculty-based VE handles the light-case and similar cases well but see it as ill-equipped to account for the subtler forms of "high-end" knowledge mentioned above. The problem is that the functioning of cognitive faculties in right environments does not seem sufficient for all cases of such knowledge (Roberts and Wood, 2014, 10-11). However, I believe this sort of VE, or to be more specific, agent-reliabilism, also accounts well for such cases. The reason is that its focus is not only on mere natural faculties but also the possibility of developing sophisticated thinking-habits and cognitive skills which meet requirements for being epistemic virtues (see Greco, 2007, 177 and further discussion below). This opens up the possibility that some forms of subtle knowledge arise from gradually acquired cognitive skills which meet conditions for being virtues. By opening up for treating skills this way, this theory differs from that of Zagzebski, which depicts virtues as character-traits that relate to the identity of the person in a deeper sense than skills. Skills, by contrast, are by Zagzebski depicted as specific, reliable procedures that virtuous persons tend to learn in their virtuously motivated endeavour to find truth in a given domain of enquiry (Zagzebski, 2008, 448). This view offers another way to capture the obtainment of subtle forms of knowledge, but it does not appear superior to faculty-based VE, whose view of gradually acquired skills as epistemic virtues also can capture "high-end knowledge" but using less restrictive requirements for what virtue is, thus avoiding the problem of restrictiveness described above. In light of this, we have grounds to prefer faculty-based VE.

This being said about necessary and sufficient conditions for knowledge and justified belief, which after all is a narrowly delimited epistemological concern, in a more general

epistemic appraisal of human agents and their performances in various intellectual domains and social contexts, the sort of virtues Zagzebski (2008) describes, such as intellectual courage and open-mindedness, clearly are relevant (see Roberts and Wood 2014). This is something we can recognize without departing from faculty-based VE. As Ernest Sosa (2017) for example argues, having morally praiseworthy character-traits can be an important aid in an attempt to find the truth, especially concerning subtle matters, even though those traits are not constitutive of the knowledge one then obtains (Sosa, 2017, 140-145). The traits Zagzebski describes might for example be involved in helping the subject carry out the right intellectual work to obtain subtle moral or scientific knowledge.

Finally, two caveats about the following investigation.

Suppose one is interested in whether *a subject is virtuous* as opposed to whether *a given trait or faculty in this subject is a virtue* (for this distinction, see Callahan and O'Connor, 2014, 8-9). Then, a broad investigation which not only looks at specific faculties or habits but also at how they relate and make up an integrated whole seems relevant. According to agent-reliabilism, a virtuous subject has an overarching disposition to employ reliable processes, methods, and habits when forming beliefs (Greco, 2007, 177). To investigate this (overarching) disposition, it seems necessary to break it down into the more specific dispositions it involves, and to find out whether each of them is reliable in the right way. Moreover, one must also investigate whether these dispositions make up an integrated whole. Greco describes virtuous character as consisting of cognitively integrated dispositions. The integration of a disposition has to do with how its output relates to the output of other dispositions. For example, a disposition whose output has no meaningful relation to the output of other dispositions is not integrated (Greco, 2010, 151-152, more on this below). This sort of general appraisal of character is not what I shall carry out with respect to theists. I shall investigate whether the theistic disposition is an epistemic virtue, not whether theists in an overall sense are virtuous subjects.

In virtue of carrying out this investigation, I do not purport to shed light on the epistemic status of theistic beliefs in *all possible cases*. As pointed out in chapter 3, section 2, it is possible to form theistic beliefs in other ways than those I have focused on. I argued that *in typical cases*, theistic beliefs arise in ways that do not involve reasoned argument or conscious reflection. Rather, they arise from the theistic disposition. I then proceeded to provide an ontogeny of this disposition, an account of how subjects become increasingly prone to form theistic beliefs in a range of situations and circumstances (see Ch. 3, section 3). The theistic disposition consists of a propensity to form largely situational, theistic beliefs without reasoned argument or reflection.

This includes what I have called "intuitive" and "explanatory" beliefs. We then saw that a propensity to form *folk-theological* beliefs can also become part of this disposition. I pointed to the frequent repetition of religious teachings characteristic of the "doctrinal mode of religiosity" (Whitehouse, 2004) and a mechanism known as "dissonance-reduction" to illustrate how this happens (section 5). This thesis about belief-formation applies, in the light of CSR, to representative cases and is therefore compatible with the view that theists in *some* cases obtain theistic beliefs in more subtle and reasoned ways, for example using acquired intellectual habits or by manifesting the praiseworthy character-traits Zagzebski, Roberts, and Wood describe. We can let it remain an open question what epistemic status theistic beliefs arising in such ways have.

Let us now look closer at agent-reliabilism, which forms the normative basis for the arguments in the next chapter.

3. Reliabilistic virtue-epistemology and the notion of an epistemic virtue

According to agent-reliabilism, when a subject knows things, then the subject gets things right as the result of the operation of an epistemic virtue. This means a correct explanation of why the subject formed true belief crucially involves the virtue (Greco, 2010, 10-12, 71-74). Justification (i.e. doxastic justification) requires that an epistemic virtue is crucially involved in the formation of the belief (Greco, 1993, 2007, 176-177).[97] The theory therefore does not (in contrast to Zagzebski) require for knowledge or justified belief that the subject manifests character-traits like intellectual courageousness or open-mindedness. What matters is that the part of cognitive character which is relevantly involved in the belief-formation meets requirements for being an epistemic virtue:

> "On the present view, knowledge and justified belief are grounded in stable and reliable cognitive character. Such character may include both a person's natural cognitive faculties and her acquired habits of thought." (Greco, 2007, 177).

[97] Greco's theory distinguishes doxastic justification (which he sometimes calls "positive epistemic status") from "subjective justification". To have a subjectively justified belief is to have a belief which seems appropriate from one's point of view. Subjective justification does not require reliability, but it is not sufficient for knowledge (see Greco, 1993, and 2007, 180). Our focus will be on doxastic justification, which on this theory is a reliabilistic notion.

Notice the two candidate-virtues here, *faculties* and acquired *thinking-habits*. In section 5, we look at how we can think of those notions in the light of cognitive science. I will there show how we can see the theistic disposition (an acquired thinking-habit) as a possible epistemic virtue. First, we will look closer at what requirements faculties and habits must meet to be epistemic virtues.

According to agent-reliabilism, both faculties and habits are dispositions to form beliefs about distinct subject matters. If these dispositions are stable, reliable, and make up integrated parts of the overall cognitive character of the agent, then they are epistemic virtues. This means that they are *abilities* (Greco, 1993, 2007, 2010). It seems fully possible on this sort of theory that some acquired thinking-habits (which might not occur widely in a population) can (if reliable in the right way) become specialized abilities that enable sophisticated scientific knowledge or subtle forms of self-knowledge. As mentioned, the theory therefore does not apply only to those forms of regular knowledge we obtain by the mere functioning of faculties. But at the same time, by not making the involvement of praiseworthy character-traits necessary for knowledge, the theory also captures regular knowledge. The crucial epistemic issue is not whether the disposition responsible for our beliefs is a natural part of the functioning of all human minds, or whether it is related to a more specific belief-forming habit that some of us acquire in particular contexts. The crucial issue is whether the disposition meets requirements for being an epistemic virtue, i.e. an ability.

Abilities in general are stable dispositions to achieve things in a relevant domain of activity or problem solving. They need not involve the formation of belief. For example, most of us have the ability to walk. Some of us can also swim, speak French, and perform well in archery. According to Greco,

> "S has an ability A(R/C) relative to environment E = *(sic.)* Across the set of relevantly close worlds W where S is in C and in E, S has a high rate of success in achieving R." (Greco, 2010, 77).

R stands for a range of tasks or challenges with respect to which one must tend to be successful in order to have ability A. C stands for conditions, meaning fluctuating or variable circumstances within environments. Greco provides examples. Whether S is sober, drunk, has sand in his eyes, or someone suddenly turns off the light, are conditions. S has the ability to hit baseballs if S tends successfully to do this in a range of environments and conditions relevant

for this ability. This is consistent with the possibility that in some environments and some conditions, S's reliability in baseball decreases dramatically. As long as these circumstances represent remote possibilities or the conditions that affect reliability negatively are not appropriate for the ability, they do not threaten S's ability (Greco, 2010, 76-77). In support of this, Greco appeals to the notion of possible worlds and argues that remote worlds are not relevant. For example, a world in which one is a brain in a vat or the victim of an evil demon is remote (assuming the actual world is the way it seems) and therefore not relevant. The fact that S is not successful in such a world is therefore consistent with S's having an ability in the actual world. Moreover, the fact that S does not perform well as a baseball player if someone turns off the light does not threaten his ability to play baseball, since such conditions are inappropriate for this ability. The important question, Greco argues, is whether S is constituted such that given this constitution, in a range of close possible worlds and when in appropriate conditions for a given ability, S tends to perform well with respect to a relevant range of tasks related to that ability (Greco, 2007, 206-215).

Epistemic virtues are examples of *cognitive abilities.* That is, they are abilities that involve the formation of belief. Such abilities are dispositions to form true and to avoid false beliefs with respect to a restricted set of relevant issues in a range of appropriate conditions. We can think of the set of relevant issues that bear on a cognitive ability as a set of propositions relevant for this ability. The important question is how S performs with respect to the propositions in that set. S has an ability A to know that p if S tends to get things right and to avoid mistakes with respect to a relevant set of propositions for A in which we can situate the proposition that p (Greco, 2007, 216). This means it is not sufficient that S just gets p right. S must also tend to get propositions related to p right. This means S's reliability must be *global* (pertain to more than one proposition) rather than local (just associated with one proposition) (for this distinction, see Goldman, 1986, 44-45).

For a disposition in S to be reliable the way virtue requires, it must involve a *tendency* to form true rather than false beliefs with respect to a restricted range of propositions, when S is in a relevant range of environments and conditions. Reliability is, as Goldman (2015) and Alston (2005) argue, a dispositional notion. Reliable processes have a tendency to provide true and to avoid false beliefs. Goldman and Alston argue further that because tendencies only apply to something that can occur more than once, reliability requires repeatability. That does not mean a reliable process actually is repeated. A process can be reliable even if it has not generated a large number of true beliefs or any beliefs. It is reliable if it *would* tend to produce true rather than false beliefs *if it were implemented.* In this way, a process can have a propensity

to generate true and to avoid false beliefs even though it has not produced beliefs. Reliability can also in some cases require that one abstains from forming belief. Imagine that the subject is in an appropriate environment and condition where there is some pressure to form a false belief. Reliability requires that one tends to withhold belief in this kind of situation. But that is not to say that occasional mistakes cannot happen. What matters is not the outcome in only a specific instance but the likely ratio of true to false in a representative sample of beliefs that come about if the process were to occur on a number of occasions (Alston, 2005, 112-119, Goldman, 2015, 37-38). Whether this ratio is good with respect to a representative sample of outputs of the theistic disposition in a range of appropriate cultural environments and conditions is a central issue in the next chapter.

To assess the epistemic status of beliefs, we need to investigate whether an epistemic virtue has been relevantly involved in their formation. This is akin to investigating the belief-formations as performances and asking whether they count as virtuous. Ernest Sosa appeals to an analogy between belief-formation and archery when discussing such an epistemic evaluation. Consider S's formation of the belief that p and an archer's shot at a target. Both performances allow evaluation with respect to success, i.e. accuracy. An accurate belief is true, and an accurate shot hits the bull's eye. We can next evaluate the performances with respect to whether they exhibit skill. Sosa calls performances that exhibit skill "adroit". Such performances might lead to success or failure. Finally, performances that are successful *because* they exhibit skill are "apt". An explanation of their success thus appeals to skill (Sosa, 2007, 22-24, 77-78, see also Greco, 2010, 74-75). Knowledge is true belief from ability. In other words, it is apt success. In contrast, a justified belief need not be successful (true), but it must arise from an epistemic ability. It is, in the current terminology, the result of an adroit performance.

The success we value the highest is the one that we best explain by pointing to stable features of the subject. The success of the skilled archer is for example to her merit in the sense that we best explain it by pointing to features she has rather than features of the environment which happened to be beneficial. Similarly, the good record reliable belief-forming dispositions usually have is also something we best explain by pointing to features of the subjects that manifest them rather than the environment in which they operate. This means we can vary environmental factors within the range of what is appropriate for these dispositions, without altering the likely ratio of true to false which accrues from their manifestation. But this does not mean that the environment is irrelevant with respect to explaining virtuous success. As Greco (2010, 140) puts it, such success always requires "cooperation from the world". To

illustrate this cooperation, he gives this example: for a courageous person to successfully save someone from a burning house, the floor must not collapse (Greco, 2010, 140). The environment thus has to be appropriate in a basic way for virtuous success to be possible. This is not the only way in which the environment is relevant. Virtuous subjects often successfully make use of resources made available by the social environment when they form beliefs. Then, a full explanation of their success includes an account of how the information was acquired through social interaction. However, for this success to count as virtuous, it must still be an inner disposition in the subject, such as an ability to assess testimony reliably, which *primarily* explains this success (Greco, 2010, 81-82). As Sosa maintains, the virtue theorist should appeal to both dispositions and situational factors when evaluating the performance of the subject. In this way, one avoids neglecting the explanatory power of the situation, and one avoids neglecting the explanatory power of dispositions (Sosa, 2017, 185). These scholars thus argue that for virtue to enable knowledge, it must be the ability (epistemic virtue) which *primarily* explains the success. This means that it is not sufficient for virtuous success that an ability *somehow* was involved. Instead, as Greco (2010, 74) argues, in a correct explanation of S's success, S's ability must be *explanatorily salient*. It must stand out among a range of relevant factors as an explanatorily crucial factor, given our interests in accounting for that success. Greco argues further that stable features of the subject are often not salient in cases where various abnormal elements have been introduced in belief-forming processes that are reliable. In such cases, the abnormal elements and not the ability stand out in an explanation of the success and this threatens the virtue of that success (Greco, 2010, 74-75). We return to this issue below.

4. Why mere reliability is insufficient for justification. True-temp and Norman-cases

We can contrast agent-reliabilism with *process-reliabilism* on the one hand and Plantinga's *proper functionalism* on the other. These three theories share the plausible view that reliability is necessary for justification (and that internal factors are not), but they offer different analyses of what role reliability plays and whether it is sufficient for justification. The present section deals with an objection to process-reliabilism. This theory states that the justification of beliefs is a product of the reliability of the process by which they arise (see e.g. Goldman, 2015). The

objection attempts to show that this is wrong since reliability is not *sufficient* for justification (see Goldman and McGrath, 2015, 40-41). Presenting this objection sets the stage for a comparison of how agent-reliabilism and proper functionalism handle this type of objection. I argue that the former does this in the most plausible fashion.

The following cases are attempts to demonstrate the insufficiency of reliability.

Case 1: The helpful demon

Greco (2007, 175-176) asks us to consider a gambler who reasons according to the gambler's fallacy. The gambler reasons that if a number has not come up for a while, then it is more likely to come up in the future. As it happens, a helpful demon adjusts reality in such a way that the gambler's reasoning tends to produce true results.

Case 2: The brain-lesion

Plantinga (cited in Greco, 2007, 175) asks us to consider a subject who has a brain-lesion which reliably triggers correct beliefs about this lesion and nothing else. The subject lacks independent evidence for the existence of this lesion and for the truth of these beliefs.

Case 3: True-temp

Keith Lehrer (cited in Lyons, 2009, 114) asks us to consider True-temp, a subject in whom someone has secretly planted a device which in a completely reliable fashion produces beliefs about the temperature. True-temp lacks evidence for or against the existence of this device and simply accepts the output without reflecting any further on it.

Case 4: Norman the clairvoyant

Laurence BonJour (2008, 368) asks us to consider Norman, a subject who has a reliable clairvoyant faculty. Norman has no evidence for or against the claim that he has this faculty or that it is possible to have such a faculty. One day, his clairvoyance produces the belief that the president is in New York City. Norman lacks independent evidence for or against this belief.

According to process-reliabilism (Goldman, 2015), reliability is sufficient for justification. This means that the subjects in these cases are justified. But this result is problematic. Intuitively, the beliefs in these cases are unjustified (perhaps with the exception of the last case, as I shall say more about below). This indicates that mere reliability is insufficient for justification. This challenges process-reliabilism. These cases point us towards agent-reliabilism or proper

functionalism, which agree that reliability is insufficient yet maintains that the reliability of dispositions or faculties is necessary for justification or warrant. But do these theories handle the cases equally well? Moreover, what other issues might be relevant in choosing between them? As we shall see, the role of a design-plan in proper functionalism is one such issue.

We start our comparison by considering how agent-reliabilism handles the above cases.

In the helpful demon case, the explanation of the success of the subject is not a reliable belief-forming disposition integrated in the subject's cognitive character. Rather, the interventions of the external agent who adjusts reality explains the success. This success is therefore not to the merit of the subject. The subject uses a process which usually is unreliable but happens to be reliable in this case because of the demon's interventions. It is then not true that in a range of close possible worlds the same process tends to yield correct results. In many such worlds, the demon does not intervene (Greco, 2010, 149-150). In light of this, epistemic virtue has not been relevantly involved. Agent-reliabilism here yields the correct result that the beliefs described are unjustified.

In the brain-lesion case, we best explain the reliability of the process by which the beliefs arise by pointing to something within the subject rather than in the environment. In other words, features of the subject rather than the environment stand out as explanatorily crucial. We can also stipulate that the lesion is a stable feature of the subject and does not pop in and out of existence. According to Greco (2010), the primary epistemic problem is that we cannot plausibly view the lesion as an integrated part of the subject's cognitive character. Integration has to do with how the outputs of a disposition relate to the outputs of other dispositions. A disposition integrated into the cognitive character of the subject has outputs that cohere with and relate meaningfully to outputs of other dispositions. Integrated dispositions "cooperate" with one another. However, this is not the case with the brain-lesion, Greco points out. It produces beliefs that have no meaningful relation to any other beliefs produced by the subject. It then cannot plausibly be considered a virtue (Greco, 2010, 150-152). The beliefs it gives rise to therefore are unjustified.

In True-temp, beliefs arise as the output of a device someone else designed and implanted. This device does not seem to be a cognitively integrated part of True-temp's epistemic agency. For example, it is not a faculty that makes up part of his natural cognitive architecture, and it is not a belief-forming habit True-temp has gradually acquired. It is an alien device secretly planted in his head. Moreover, given the presence of this device, the process of spontaneously forming beliefs about the temperature without having any relevant evidence is reliable. But given the absence of this device, the same process would be unreliable. In a correct

explanation of True-temp's epistemic success, the agency of the designer of the device rather than features of True-temp therefore stands out as the primary, salient factor. This success is therefore to the merit of the designer rather than True-temp and hence, it is not virtuous. The designer rather than True-temp is the one who achieves something when True-temp forms true beliefs about the temperature. In light of this, the beliefs the device produces are not justified.

Norman's epistemic success is not the result of external intervention. For example, no one tampers with him by implanting a clairvoyance-device in his head or by secretly adjusting external reality so that his beliefs about the president always turn out correct. Furthermore, the clairvoyance is not a disorder akin to the brain-lesion. This reliable disposition, which possibly is integrated into Norman's overarching agency, is responsible for the success. Could we think of the clairvoyance as a reliable habit which Norman has (somehow) acquired, and which has become (or might become) integrated into his overarching agency? Possibly. But then, agent-reliabilism does not obviously rule that the beliefs are unjustified. I find this unproblematic because it is not intuitively obvious (at least to me) that the beliefs in fact are unjustified.

Agent-reliabilism thus leads to plausible results in cases that indicate that mere reliability is insufficient for justification. The theory achieves this by restricting what kind of reliable processes that can confer justification. Only reliable processes that have their basis in the functioning of cognitive abilities do this. The abilities must be reliable in the right way, and abilities can be ascribed to the agent who then can take credit for the success they explain (Greco, 2007, 2010).

How does Plantinga's *proper functionalism* handle these and similar cases? By scrutinizing this, we can appreciate where this theory converges with agent-reliabilism and where the two diverge. This helps us see why agent-reliabilism is more plausible. Recall that according to proper functionalism (see chapter 5, section 2), warranted beliefs arise from faculties that are functioning properly and reliably in an appropriate environment in accordance with a design plan aimed at truth (Plantinga, 1993, 19). Then, clearly, mere reliability is insufficient for warrant. This theory provides plausible verdicts in the two first cases. Consider the gambler who gets things right as the result of the benevolent demon. Given the presence of the demon, he is not in an appropriate environment, that is, the kind of environment his cognitive faculties were designed to operate in. His beliefs therefore lack warrant even if true and reliably produced. The beliefs of the lesion-victim also lack warrant. The lesion is a pathology, something which violates the proper functioning requirement. Here the verdicts agree with those we provided above. As we will see below, the True-temp case proves more difficult. Here agent-reliablism and proper functionalism depart from each other because the

latter seems to implausibly grant warrant. The reason for this has to do with the notion of a *design-plan* in Plantinga's theory. Let us briefly consider this plan.

A design plan is a set of specifications for how something is supposed to function (Plantinga, 1993, 14). Plantinga points out that our cognitive faculties might be designed to achieve more than just the production of true beliefs. A full design plan contains specifications that indicate all the purposes they have. For example, parts of their purpose could be to comfort us or help us achieve other goals, perhaps through exaggerated optimism. Plantinga stresses that *only* those beliefs that arise in accordance with that part of the design plan which is *aimed at truth* have warrant. This means that a belief arising from a properly functioning faculty in an appropriate environment and which is reliable in that environment but does not produce beliefs in accordance with a truth-aimed design plan is not warranted (Plantinga 1993, 15-16). This is too restrictive. Here proper functionalism and agent-reliabilism diverge in ways that provide grounds to favour the latter. Before appreciating why, let us turn to the True-temp case and an additional similar case. Here the theory is too *permissive* in granting warrant.

As Jack Lyons points out, the secretly implanted device in True-temp presumably has a design-plan that specifies that generating true temperature-beliefs is its constitutive aim. It is then not a device designed to do something else, but which happens to accidentally produce correct temperature-beliefs in addition. Rather, someone has deliberately designed it to get things right with respect to the temperature. Furthermore, the device is (by stipulation) reliable. It achieves its purpose. Finally, it is operating in the right kind of environment, namely True-temp's head (Lyons, 2009, 128). Proper functionalism therefore rules that True-temp's beliefs have warrant, but this is problematic (as I argued above).

Plantinga gives us the next example. He asks us to suppose that God created us in such a fashion that religious beliefs arise from wish-fulfilment, and he contends that this is compatible with warrant. The reason is that the beliefs would in such a case arise from a properly functioning faculty which is reliable in the right kind of environment according to a design-plan aimed at truth (Plantinga, 2000, 197-198). Let us then imagine that God has provided our minds with a wish-fulfilment device which is responsible for our religious beliefs. Is such a device a cognitively integrated system or is it more akin to the device implanted into True-temp? In the current case it seems that God has not tampered with our minds the way the designer of that device did. That is, he has not inserted an alien device into our minds but designed our minds such that they already contain a wishful thinking device. Suppose it therefore is cognitively integrated. Then, a cognitive integration-requirement is met. Does it follow that the faculty produces warranted beliefs? A challenge here arises. It seems, as was the

case with True-temp, that we do not best explain our present epistemic success using this faculty by pointing to a feature we have. The primary factor which stands out is God, the supernatural external agent who designed us with a weird yet reliable *wish-fulfilment* system. The presence of God in explaining our success from this system challenges warrant. Let us briefly explore how.

We start by observing that the problem is *not* that God, a supernatural agent, is part of an ultimate explanation of this faculty. A theist might for example contend that God is the ultimate explanation of all of our faculties since God guided evolution to lead to their formation. God's role would then be analogous to the role of evolution by natural selection, which to naturalists is the ultimate explanation of why we have the cognitive makeup we have. It is not plausible that the role of God or naturalistic evolution in an ultimate explanation would hinder warrant.[98] The primary problem for the wish-fulfilment system arises from how God is a *salient* part of a best proximate explanation of our success from this faculty. To see this, we need to focus on what is salient in this case. Greco (2010) argues that an explanatorily salient factor need not be the only relevant factor in accounting for some event. For example, in explaining a car crash where the driver was drunk, the drunkenness is the salient factor but other factors such as the weather, the quality of the road or the tires, and so on, might also have contributed to the crash (Greco, 2010, 74). Greco argues further that salience usually involves a deviation from what is normal. For example, in explaining how a fire broke out in a welding facility, flying sparks are not salient even though they might have contributed to the fire. By contrast, the unusual presence of flammable chemicals immediately stands out as the salient factor (Greco, 2010, 75). With this in mind, let us next observe that the process of forming beliefs by wishful thinking is in general unreliable. The study of wish-fulfilment and its role in the formation of religious beliefs will therefore not shed much light on why *religious* wishful thinking is an epistemic success. That would be like trying to explain the fire by pointing only to the sparks or explaining the crash by pointing to the state of the car. What we need, instead, is to find the analogue of the drunkenness or the flammable chemicals. This brings us to God. To explain why religious wish fulfilment is reliable despite the fact that wishful thinking usually is unreliable, we must appeal to God. For example, God might have rigged the wish-fulfilment faculty so that it never leads us to wish for religious things that are not true (for example, it would never trigger the wish that there are demons in heaven) or perhaps he continually adjusts

[98] Plantinga (1993) does not agree, since he rejects the thesis that our beliefs have warrant if the underlying faculties arose through naturalistic evolution. This relates to the restrictiveness of proper functionalism, which we discuss below.

things in the supernatural realm in ways that ensure that whatever we wish for concerning religious issues turns out true. The central point is that unless we appeal to God, the success of this otherwise unreliable process would be a puzzle. In a correct explanation of why religious but not other forms of wish-fulfilment yields correct beliefs, the element of supernatural design and supernatural intervention thus stands out as salient in a way the psychology of wish-fulfilment does not, even though that is also a relevant part of the explanation. In conclusion, given that we do not best explain our success by pointing to features we have, this success it not virtuous and the beliefs not justified. This, together with the rulings on the True-temp case, indicates that Plantinga's theory is too permissive.

The notion of a truth-aimed design plan plays an important role in Plantinga's assessments of the above cases. To notice this, simply remove this plan from True-temp's device or the religious wish-fulfilment device. Suppose the designer of these devices designed them for other purposes but that their functioning nevertheless led to epistemically successful outcomes with respect to the temperature or religious issues. Then, the theory yields the result that the beliefs lack warrant. But this is not plausibly why they would have that epistemic status. As my analysis indicated, a truth-aimed design plan or lack thereof does not impact on the warrant of these beliefs. They would lack warrant also with such a plan (for the above-indicated reasons). This brings us to some respects in which proper functionalism is too *restrictive*. Two considerations show this. First, the theory seems to be unacceptable for those not ready to embrace theism. It is therefore too restrictive in the sense that it is committed to a specific religious worldview many will not accept. Second, the theory fails to grant warrant where truth-aimed design-plans are absent but we have good reasons to grant warrant nevertheless.

We should appreciate how Plantinga links the warrant of beliefs to the truth of theism. He does this in virtue of making it a requirement for warrant that beliefs arise from the proper functioning of faculties in accordance with truth-aimed design-plans, and then arguing that we cannot account for the latter ideas from a non-theistic perspective. As Plantinga explicitly points out, the epistemology he promotes can *only* flourish given "supernaturalism" (1993, 194). The reason is that its crucial notions of proper functioning and truth-aimed design plans according to Plantinga pertain primarily (and perhaps only) to consciously designed artefacts. While (metaphysical) naturalism does not allow us to see human beings (and human cognitive faculties) in such a fashion, supernaturalism (i.e. theism) allows this. The theist can see human beings as well as things like airplanes or coffee-machines as consciously designed and therefore as artefacts. Then (and perhaps only then) we can also readily apply the notions of proper functioning and design-plans to human cognitive functioning (Plantinga, 1993, 195-198).

Furthermore, Plantinga argues, central attempts by naturalists to account for the notion of proper functioning in general (as applied even to things like hearts, kidneys, and so on) fail. Given naturalism this "looks impossible" (Plantinga, 1993, 211). These strong claims illustrate how Plantinga weds his epistemology to theism in a way that leads to a highly restrictive theory. If it really is true that we need to assume theism to make full sense of its central ideas, then perhaps we should simply reject those ideas and pursue another theory. Plantinga would perhaps respond that it is no objection to the correctness of his epistemology that we cannot account for its central terms or plausibly see our minds as meeting necessary requirements for warrant unless theism is true. To the extent that we find his theory correct (and therefore accept its central notions and requirements) we should instead conclude that *theism* is true.[99] But from a non-theistic perspective, one might instead simply reject a theory which is wedded to theism in such a fashion. As a basic requirement, one might suggest, an epistemological theory should be able to make sense of how we can justifiably believe or know things without presupposing that a specific (and strongly contended) religious worldview is correct. Proper functionalism is problematic in how it violates that requirement.

The second consideration concerns the restrictiveness involved in making the presence of truth-aiming design-plans necessary for warrant. We have strong grounds to believe that the ultimate function of our cognitive faculties is to contribute to fitness. This means that if these faculties have something like design-plans, then these plans probably specify that the ultimate goal of our cognition is to help us survive and reproduce. According to Plantinga, this makes it "doubtful" that our faculties (in terms of proximate functioning) are reliable (1993, 218). But this seems wrong. We have grounds to see faculties that aim (ultimately) at fitness to be largely reliable with respect to tasks for which they are specialized. Wilkins and Griffiths (2012) plausibly contend that reliability – in the sense of relatively accurate tracking of relevant aspects of our surroundings – *is* a central way our cognition contributes to fitness. Reliability is therefore something we can expect from our natural cognitive faculties, and it is wrong to think that faculties aim either at fitness or reliability as if these were mutually exclusive options (Wilkins and Griffiths, 2012, 134-135, see also Barrett, 2015). Here is how the restrictive nature of proper functionalism reveals itself. Suppose Plantinga's doubts about reliability are

[99] As Plantinga suggests, those not ready to discard the idea of proper functioning but who at the same time cannot see how we can make sense of it unless theism is true have the basis for an argument against (metaphysical) naturalism and for theism (1993, 214). We should be aware that Plantinga also provides an epistemic argument, which aims to show that one cannot rationally accept (metaphysical) naturalism because it is a self-defeating position (Plantinga, 1993, 216-237). This is an example of an evolutionary debunking argument (EDA), given how it uses evolutionary considerations to undermine the idea that our cognitive faculties are reliable. We will not in the present investigation discuss this complex and thought-provoking argument.

unwarranted and that our cognitive faculties are reliable despite ultimately aiming at fitness. On proper functionalism, their functioning does not then enable us to justifiably believe or know things. The reason is that they lack the right kind of design-plan. This is problematic. It is too restrictive to claim that cognitive faculties operating reliably in their right kind of environment cannot enable us to have justified belief or knowledge simply because their ultimate aim is not to give rise to true belief. Agent-reliabilism here is more plausible. We saw that on this theory, it suffices for a faculty to be an epistemic virtue that it is reliable in the right way. This is a more plausible theory, which allows us to see how we can know and justifiably believe things even though a naturalistic explanation of how our cognition arose is correct.

In summary, and to connect some threads, the argument against internalism in chapter 4 placed us in a general externalist landscape. I then in chapter 5 criticized a central example of religious, externalist epistemology, namely Alvin Plantinga's theory. The present chapter first briefly criticized Linda Zagzebski's version of VE on the grounds that making virtuous motivation and the presence of virtuous character-traits necessary for virtuous performance (and in extension, justification) is too restrictive. I then defended agent-reliabilism and argued that it is more plausible than proper functionalism: Agent-reliabilism has an advantage to proper functionalism in virtue of how it handles True-temp cases and in virtue of how the latter implausibly makes the presence of truth-aimed design-plans necessary for warrant. This brings us to the issue of what epistemic status the theistic disposition and by extension, the beliefs that arise from it have. As our discussion in chapter 3, section 3 showed, in typical cases, subjects obtain this disposition gradually as a by-product of frequent intuitive thinking about God. We now ask if this disposition is an epistemic virtue.

5. The theistic disposition as a potential epistemic virtue

Before pursuing this issue in detail in the next chapter, let us first briefly clarify what sort of candidate virtue this disposition is. We saw above that from the perspective of agent-reliabilism, both cognitive faculties and acquired thinking-habits are potential epistemic virtues. How can we best envisage their nature? The theory is not wedded to any particular psychological theory of cognitive faculties, but in the light of CSR, which is our current focus, we can interpret them as domain-specific, functional systems. This is for example how we interpreted Plantinga's purported god-faculty (Ch. 5, section 2), and it is the way we interpreted the agency-detector

(Ch. 2, section 3), the attachment-system (Ch. 2, section 7), and the cheater-detector (Ch. 3, section 3). When it comes to belief-forming habits, agent-reliabilism is open about what specific structure they have or what kind of cognitive operations they carry out. As examples of habits which potentially are virtues, Greco mentions skills subjects develop through training and exercise (2007, 177). We can envisage sophisticated habits involving complex intellectual procedures or less sophisticated habits governed more by intuitive thinking. What matters epistemically is that such habits are reliable with respect to a correctly restricted range of tasks, and that they are cognitively integrated into the agency of the subject (Greco, 2007, 211).

How does the theistic disposition fit into these categories? As a start, it is not a faculty but a gradually acquired habit involving the operation of numerous faculties. Moreover, given its ontogeny and functioning (Ch. 3, sections 3, 4, and 5), it is not an intellectual habit involving a propensity to engage in complex, reasoned procedures. It involves mostly intuitive thought and little explicit reasoning. Before we analyse it in detail in the next chapter, what more can we say about it in general terms? One might for example wonder whether it involves what Robert McCauley (2013) calls "practiced naturalness". By this, McCauley means a sort of tacit expertise whereby the subject employs internalized and habituated procedures that require cultural input, instruction, and practice to bring about. As examples, McCauley mentions the tacit expertise a chess-master has, or the abilities to ride a bike, read, or write (McCauley, 2013, 4-5, 24-25). This clearly relates to what Greco (2007) refers to as skills developed through training over time. For example, mastery in riding a bike is an ability and this ability involves practiced naturalness. We cannot yet determine whether the theistic disposition is an ability (i.e. epistemic virtue). That is what the next chapter will do. But it seems fair to say that the theistic disposition involves practiced naturalness, at least in the following sense: it arises gradually as the result of a *repetition* of relevance-enhancing episodes whereby an intuitive, *culturally specific* concept is activated. Moreover, God, who to non-theists does not seem real, seems real. We have also seen (Ch 3, section 5) that evangelical Christians learn to employ particular methods to develop an even more direct experience of being personally related to God (Luhrmann, 2012). To the extent that mastery of such methods is integrated into theistic dispositions, such dispositions also involve this additional practiced naturalness. One might then say that in a "practiced" sense of naturalness, the (assumed) reality of God has become natural, sometimes in a very strong sense whereby God is felt to be part of mundane aspects of one's daily life.

Given the centrality of *intuitive,* automatic cognition in the theistic disposition, does it also involve something like what McCauley (2013) calls "maturational naturalness"? As we

saw in chapter 5, cognition which is natural in this way is produced spontaneously, automatically, and without the need for specific cultural input or instruction (McCauley, 2013, 5). Since the theistic disposition is culturally specific, it is not natural in this sense. Yet, it has an element of naturalness in the maturational sense because the underlying cognitive processes that enable it to arise and thrive - when the mind is exposed to cognitively attractive input – are natural in that sense. As we saw in chapters 2 and 3, the way the mind constructs the concept of God and the way the relevance of this concept subsequently is enhanced and preserved in the mind strongly depend on automatic and intuitive cognition produced by evolved cognitive systems. In light of this, maturationally natural cognition is present in the theistic disposition. This disposition therefore does not straightforwardly fall into either of the two categories McCauley discusses.

In terms of the way these categories carve our cognition, the theistic disposition seems to be a hybrid belief-forming habit with elements of both practiced and maturational naturalness. There is support for this notion that religious belief-formation and activity combines these forms of naturalness. Helen De Cruz and Johan De Smedt (2015, 37-38) for example argue that religious beliefs and practices often involve a combination of these two forms of naturalness rather than mostly maturational naturalness. To illustrate their view, De Cruz and De Smedt provide the striking example of *glossolalia*, speaking in tongues. This practice exploits our (maturational) natural language-capacity, but it is also dependent on the right cultural input, and some level of practice (De Cruz and De Smedt, 2015, 37). As T. M. Luhrmann points out, evangelical Christians interpret glossolalia as a divine gift, but from a psychological point of view, it is better construed as a skill subjects gradually learn to master and which they in turn can teach others how to master (Luhrmann, 2012, 24-25). As another example of how practice can be combined with cognitively natural processes, Luhrmann describes the particular sort of *prayer* subjects learn to practice in these communities as a skill the use of which can help make God appear more real (2012, 133-134). Finally, that religious beliefs and practices combine maturational and practiced naturalness is also a theme we saw that Barrett and Lanman (2008) stress in their discussion of how some culturally specific practices can "stretch" beliefs some distance away from the attractive structure our non-conscious "maturational" cognition favours, but not too far away (see also McCauley, 2013, 207). We saw (Ch. 3) that such practices cannot stretch our intuitive or habitual thought too far away from maturational naturalness, given how it acts as a constraining factor. This is why formation of beliefs that get complicated theology (such as the doctrine of the Trinity) right is unlikely to become habitual. By contrast, beliefs that get such theology wrong or gets easy to

grasp theology right can achieve this. The formation of such *folk-theological beliefs* can be integrated into the theistic disposition through mechanisms we have discussed (Ch. 3, section 5).

We are now ready to investigate whether the theistic disposition - a complex habit which comprises a propensity to form intuitive, explanatory, and folk-theological beliefs - meets the requirements for being an epistemic virtue. To be more specific, we shall investigate whether this disposition meets necessary requirements for being a virtue, the foremost requirement being reliability (as specified above). Given that the epistemic status of beliefs hinges on whether they arise from an epistemic virtue, our investigation can shed light on the epistemic status of theistic beliefs.

Chapter 7. The Reliability that Virtue Requires. Evaluating the Theistic Disposition

1. Introduction and chapter-outline

The theistic disposition (TD) is a habit whereby theistic beliefs arise as the result of how theistic concepts entangle themselves with various functionally distinct processes. In contrast to what Plantinga (1993) argues concerning warrant, the TD is a potential virtue even though it is not governed by a truth-aimed plan. Moreover, in contrast to what Zagzebski (2008) argues concerning the nature of epistemic virtue, the virtue of the TD does not require it to be related to virtuous character-traits. I have argued that such requirements are too restrictive. Instead, from the perspective of agent-reliabilism it is sufficient for the virtue of this disposition that it is reliable in the right way.

The present chapter investigates the epistemic status of the TD and by extension the beliefs that arises from it. We shall look at this habit from different angles, starting in section 2 with the functional processes the theistic concept latches on to. We here deal with what is known as the "generality-problem". This is an alleged problem for reliabilism in general which potentially creates trouble for reliablistic evaluations of theistic beliefs. Helen De Cruz and Johan De Smedt point to this problem in a discussion of reliabilistic arguments against the justification of theistic beliefs (De Cruz and De Smedt, 2015, 190-194). I respond by showing how an approach developed by William P. Alston known as "psychological realism" opens up for solving the generality-problem. CSR contributes to this by enabling us to specify functional procedures which point us to processes by which theistic beliefs arise. This provides additional grounds to reject P3 in E-ENT, which claims (among other things) that CSR does *not* explain by what processes theistic beliefs arise. Introducing this solution to the generality-problem for *processes* thus relevantly connects with our discussion of E-ENT in chapters 1 and 3. It also facilitates our virtue-epistemological investigation of the TD. This solution to the generality-problem facilitates the solution of a generality-problem that arises for the TD as a *complex belief-forming habit* in its own right. It does this by helping us specify the relevant type of habit

it is a token of and by restricting the range of propositions this type must tend to get right in order to be a virtue. I argue that the TD is a token of a broader type of disposition which leads to beliefs in different supernatural agents in different cultural contexts as the result of the same underlying, psychological processes. I call this broader type the "MCI disposition". With the specification of this type we are ready to assess whether the TD is reliable the way virtue requires. Section 4 presents *the argument against epistemic virtue*, which shows that the TD is not virtuous because its relevant underlying type, the MCI disposition, fails to meet the reliability-condition for virtue. The *argument against epistemic justification* then shows that the theistic beliefs that arise from the TD therefore lack justification. This negative conclusion applies to the three types of theistic beliefs specified in chapter 3, intuitive, explanatory, and folk-theological beliefs. I purport to show that this conclusion is correct even on the basis of an assumption which is compatible with theism and which theists in many cases will agree with, namely that most MCI agents do not exist. Given how CSR contributes to this evaluation of theistic beliefs, the field has normative implications for theistic religion in sense (3).

2. The generality problem and how theistic beliefs arise

We now start the evaluation of the TD by zooming in on the processes it involves. We thus respond further to the claim (criticized in chapter 3) that CSR cannot help us pinpoint the processes by which theistic beliefs arise. I shall here employ some resources in epistemology to offer additional criticism of this, thus also preparing the ground for our evaluation of the TD in section 3 and 4.

Alvin Goldman describes processes as procedures that respond to input, manipulates it according to some rule, and produce beliefs as output. Based on this, we can describe a process by plotting information into the following scheme: Input → Functional procedure → Output. As a start, Goldman describes (as examples of processes) some folk-categories that carve cognition into coarse types, such as "perception", "reasoning" and "emotion-based" processes (Goldman, 2015, 38-39). These broad and general types match categories that intuitively come to mind when we ponder various kinds of cognition, but it is not plausible that these provide the right types. In the case of perception, a strategy to move from folk-categories to something more specific is to appeal to the sensory modalities. This gives us types like "olfactory perception", "auditory perception", and "visual perception". These types also seem to reflect how human

minds intuitively categorize large chunks of cognition, and they provide a convenient way to arrange material in a textbook in cognitive psychology (for example). Textbooks in psychology also describe many fine-grained processes under each of these broad headings. However, no matter how specific we get, it is far from clear that we have correctly specified the relevant type to which a given token-process belongs. Goldman points out how critics of reliabilism allege, based on these considerations, that there are principled problems in attempts to specify the right types (Goldman, 1986, 49). Let us now consider the alleged problem reliabilism here faces and how we can handle it.

Earl Conee and Richard Feldman (2011) point out how endless the possibilities are when it comes to specifying a type for a process and restricting its range. For example, visual perception at night seems quite different from visual perception during the day. Should we construct two different process-types in order to account for this difference, or should we instead describe one process under different conditions? What then about visual perception at night when on drugs or when on drugs while walking? There seems to be no end to how many types we can construct and with what level of generality we can specify them. Conee and Feldman (2011) contend that reliabilism lacks a non-arbitrary method for specifying the right type among these options and therefore is in trouble as an account of justification and knowledge. This is the *generality problem*. This problem has two parts. The first and primary part is the alleged impossibility of correctly specifying the relevant type (Conee and Feldman, 2011, 137). We can call this *the relevant type* problem. The other part is the challenge of specifying a range of situations where the process occurs. This is crucial for process-reliabilism, since it holds the view that the justification of beliefs depends on how reliable the responsible processes are. What matters to justification in this theory is how reliable the process is when operating in an appropriate range of situations rather than in any situation (Conee and Feldman, 2011, 135-138). As we shall see in section 3, a similar challenge arises for agent-reliabilism, which must restrict the range of circumstances, conditions, and propositions that pertain to a given habit or disposition. We can call this part of the generality-problem the *relevant range problem*.

William P. Alston (2005) argues that we can solve the generality problem. This solution is, in my opinion, on the right track in how it replaces an arbitrary search for types with a method where an appeal to science facilitates a targeted non-arbitrary search. Alston suggests that among the many possible types that we can imagine, some types are special. They have this status in virtue of carving out and correctly reflecting the outline of something psychologically real. Psychology helps us pinpoint what is psychologically real and therefore helps us find the right type. We will in the present section see how this solution, together with input from CSR,

helps us identify processes by which theistic beliefs arise. This allows us to appreciate how the TD looks from an angle where we focus on functional procedures. Moreover, in section 3, I argue that this approach and input from CSR help us handle a similar generality-challenge for the TD itself, thus preparing us for conducting a virtue-epistemological analysis of this habit.

The right strategy is to look at what *functional procedure* a process involves when that process produces a belief. Alston here appeals to Goldman's definition of a process as a *procedure* that takes doxastic or non-doxastic input, manipulates it, and produces a belief. The middle part in this scheme is the procedure, which means a rule or algorithm by which a process generates a doxastic state based on input. Alston proposes that when the subject forms a belief, then this event realizes a functional procedure. He calls psychologically realized functions "psychological mechanisms". The solution to the (relevant type part of) the generality problem is according to Alston that the functional procedure that governs a given belief-forming episode provides the right type for the process involved. Alston contends that the type we obtain when we correctly identify the procedure reflects the contours of the causally relevant aspects of the belief-formation. Such a type accounts for the causally relevant psychological events that occur when the subject forms a belief. These events are unrepeatable, but the underlying functional procedure is not. This procedure is a general rule that minds can realize repeatedly (Alston, 2005, 125-130, see also Goldman, 1986, 50). Notice that cognitive science is probably not able to detect the exact, unrepeatable cognitive events that produced a belief in a given subject in a given situation. However, such science seems capable of finding the general rule involved when the mind of the subject produced that belief. Then it is, given Alston's approach, capable of finding the type of which that belief-forming episode is a token. To obtain a sufficient explanation of such a token we must in addition make some relevant observations that provide grounds to relate it to the general process we have found. How can the process be detected?

Theistic beliefs arise from processes where such a rule is identifiable. CSR helps us find this rule and therefore can help explain how the beliefs arise. As we have seen many times, scholars in CSR appeal to functional adaptive procedures, which are implemented by evolved cognitive faculties. Such faculties solve adaptive problems that they are specialized to handle. The result is output in the form of physiological reactions, motivations to act, emotions and beliefs. Such output tends, in relevant environments, to contribute to solving the adaptive problem (see for example Buss, 2015, 45-46). We can plot the relevant procedure into the following scheme in order to obtain the relevant type: 'Input → Functional procedure → Output'. The following illustrations show how we can use this strategy to detect three types of process that all lead to occasional theistic by-product beliefs.

The agency-detector carries out a functional procedure we can simply call agency-detection. This procedure then provides us the type, which we also can call agency-detection. This type of process is involved when we form beliefs about mundane natural agents. The same process sometimes also produces theistic beliefs. This happens when a theistic concept entangles itself with the functioning of the agency-detector. The process is the same even though the output is not the kind of belief that in relevant environments contributes to fitness. Consider a subject S with the TD, who has an ambiguous experience during a hike in the mountain. Various thoughts about God and a sudden gust of wind trigger the agency-detector, which signals the presence of an agent. The theistic concept latches on to these signals. As a result, the belief that God is present suddenly arises. We can plot this into the scheme:

> (Input) Ambiguous experience, cues that satisfy conditions triggering agency-detection, theistic concept → (functional procedure) agency-detection → (Output) S's belief that God is present.

In the present episode, agency-detection operates on a pattern of conceptual and perceptual input and this process then leads to an output belief about God. If we assume that S has a theistic disposition, then we can expect that many occurrences of agency-detection could yield theistic beliefs in this and similar situations. This is because in such a subject, a theistic concept has significant psychological relevance. The threshold for its activation in situations where an evolved system carries out a functional procedure is therefore low. The result would be that the subject frequently thought God were present. In subjects less prone to form theistic beliefs, the same process might occasionally produce theistic output or no such beliefs at all.

Beliefs about the likelihood that an agent with whom one engages in social exchange will defect arise from the functioning of the cheater-detector. Chapter 3 described a case where two subjects, S and Z, engaged in social exchange. A theistic concept latched on to intuitions produced by S's cheater-detector with the result that a theistic belief arose, according to which God disapproves of Z. We can plot this case into the present scheme thus:

> (Input) Information relevant to successful handling of social exchange, theistic concept → (functional procedure) cheater detection → (output) S's belief that God disapproves of Z.

Identifying Z as a defector is likely to trigger a feeling of resentment towards Z. When a theistic concept infects this process, the result becomes a theistic overlay on those reactions, something

that induces S to believe that God resents Z. This belief is an occasional by-product of a process the cheater-detector implements, in minds that already contain a theistic concept.

Attachment-theory of religion describes how God functions as an attachment-figure. God becomes such a figure when intuitive thinking tied to the (adaptive) goal of facilitating attachment to caregivers gradually extends in scope such that it eventually applies to other persons in general and finally also to a purported deity the concept of which the mind has. A psychological process that functionally is about regulating relationships to attachment-figures now can produce occasional by-product beliefs about God. For this to happen, the theistic concept must entangle itself in the process:

> (Input) a feeling in S of God being psychologically remote, theistic concept → (functional procedure) regulating relationship to attachment-figure → (Output) S's belief that God does not approve of him.

When God is an attachment-figure, then the attachment-system monitors the level of psychological proximity ("felt closeness") there is between the subject and God. Signals that indicate that this level is less than optimal, such as a feeling of God being remote, can trigger theistic beliefs whose contents strongly reflect the workings of the attachment-system. In the present example, the subject has insecure attachment-style. The content of the theistic belief reflects this. We saw (chapter 2, section 7) that insecure attachment involves a working model consisting of negative expectations towards others. In this case, those expectations extend to God.[100] Because S expects others in general not to approve of him, he also expects God not to approve of him.

We have now seen three processes by which intuitive theistic beliefs arise. By identifying the functional procedures involved, we identified their right types. An important basis for this identification-procedure is the contention that religious beliefs are by-products of how the mind solves fitness-relevant tasks unrelated to religion. We now turn to the other part of the generality-problem discussed above, namely the challenge of limiting the range for a process. This provides the basis for our challenge below of limiting the range for the theistic disposition, which is not one process but a propensity to form beliefs using various processes.

What range do the three processes we identified have? Goldman (1986) points out two dangers that an attempt correctly to restrict the range of a process involves, namely that we limit

[100] The subject might of course engage in an analytical reasoning-process that deduces that God, who by definition is perfectly good, loves him. However, such reflective reasoning is not likely to offset the power of intuitive, non-reflective thinking. We saw an example of this in chapter 5, section 3, where we looked at how a tendency to anthropomorphize God prevailed in subjects that when asked about it reported that God is omnipotent.

the range too narrowly or too broadly. If we do the former, then we encounter the "single case problem". For example, if we limit the range for a type to just one unique belief-forming episode, then the process is completely reliable if the belief it gives rise to is true. Process-reliablism then leads to the problematic result that this belief is justified simply in virtue of being true. If the episode leads to a false belief, then the process is completely unreliable and the belief unjustified, which also is a problematic result. In contrast, if we broaden the scope for a process too much, then we encounter the "no distinction" problem. Let us say we assume some general measure of reliability for a very broad type. Then the epistemic status of each belief we assume to be a token of this type will be the same, something which is problematical (Goldman, 1986, 49-50, Alston, 2005, 118).

Alston's notion of reliability as a dispositional notion allows us to avoid the single case problem. The reliability of a process does not depend on how many true beliefs it produces, but it depends on how well the process would perform if it occurred in a representative range of cases (Alston, 2005, 119-120). The types we specified do not have only a single occurrence. The functional procedures that evolved cognitive systems implement are part of strategies whereby the mind solves distinct recurrent adaptive challenges. Clearly, the types we obtained by identifying these procedures therefore have more than one occurrence. The relevant challenge is therefore the no distinction problem.

How can we avoid attaching an inappropriately wide scope to a process-type? Moreover, what exactly is it that we must restrict? The situations in which subjects find themselves seem relevant. For example, we should keep out evil demon scenarios and other bizarre arrangements where the subjects are radically deceived about everything but have no way of finding out about their plight. For example, the visual perception of a victim of an evil demon leads to false beliefs. Such failures should not affect a measure of the reliability of visual perception in mundane situations where one is walking down the street, even though the same type of process occurs in both cases. We should therefore limit the range of situations in such a way that we rule out evil demon scenarios as irrelevant. As Alston suggests, we should restrict the range to situations we "typically encounter" (Alston, 2005, 124). However, such a restriction does not seem sufficient because even though we (presumably) do not suddenly enter evil demon scenarios, it is far from clear how best to characterize the situations we typically encounter. What kind of situations are these? Moreover, different kinds of situations are probably relevant to processes that different evolved systems carry out. The important question when it comes to processes is therefore not only what situations "we" (as subjects) tend to encounter, but what sort of input various cognitive systems (i.e. cognitive faculties) within us are specialized to

handle. We therefore should focus on the following as a relevant factor that should be restricted as part of the solution to the relevant range problem: the spectrum of *information* which we allow to bear on a cognitive system. We can further relate information to propositions. That is, related to the spectrum of information that we allow to bear on a cognitive system is a set of propositions about that information. The reliability of a process can then be thought of as involving a tendency to get things right with respect to the set of propositions tied to the information that we allow to bear on the underlying system.

The world contains all kinds of information. Some of that information is appropriate input for a given type of cognitive process but not all information out there is. Information which is inappropriate for a given process might be appropriate for another process, or it might be inappropriate for human cognition in general and therefore fail to constitute input for any process. Consider processes that involve visual perception (in humans) and radiation in those parts of the electromagnetic spectrum that are invisible to us. This radiation is outside of the range of input for human visual perception. A proper measure of the reliability of visual perception need not take inability to detect this sort of radiation into account, because this is outside a plausible range of information relevant to this sort of process. In terms of propositions, an inability to get things right with respect to propositions about the radiation which is invisible to us does not bear negatively in a measure of our visual perception. For example, unreliability with respect to gamma rays does not plausibly reduce the reliability of the process that allows me to see the cup of coffee on my table. This consideration indicates that we should limit the range of information that constitutes input for a given process-type to what is humanly possible to detect and process.

A consideration that further indicates how we should limit the range is that cognitive faculties (or systems) are *domain-specific*. This means that they are specialized to deal with a specific range of information that is relevant to handling a distinct set of challenges and tasks (Hirschfeld and Gelman, 1994, 3, Barrett, 2015). A domain consists of intuitions and expectations that allow an identification and interpretation of distinct phenomena, exposure to which constitutes a recurrent challenge for the organism (Hirschfeld and Gelman, 1994, 21, Barrett, 2015, 26-28).[101] For example, the naïve ontologies we discussed in chapter 2 seem to be examples of domains in this sense, and the cognitive faculties we frequently have discussed

[101] I have amended this definition by replacing "body of knowledge" with "intuitions and expectations". This is because I use knowledge in a specific technical sense where knowledge entails true belief, and what Hirschfeld and Gelman describe with this term does not entail true belief. Domains reflect how humans intuitively carve reality, and this might not correspond with the actual nature of reality (see Ch. 2, section 4 for more).

seem domain-specific in the above sense. Distinct cognitive domains arise and stabilize in subjects as the result of the functioning of evolved cognitive systems. These systems evolved as solutions to distinct adaptive challenges, the solution to which involves a dedicated procedure (Barrett, 2015, Buss, 2015).

Information which indicates aspects of reality relevant to handling the tasks or challenges the system is specialized to handle, and information which seems like the former kind of information but without actually indicating those aspects of reality, are within the domain of a cognitive system. In evolutionary terms, both fitness-relevant information and information that just seems fitness-relevant can activate a system. The system makes "bets" when responding to such information, and the making of these bets is part of the strategy by which the system solves the problem. Such bets tend to pay off in environments that resemble those environments in which the challenge has been a recurring feature (Barrett, 2015, 23-24). Dan Sperber's distinction between *proper* and *actual* domains, which relates closely to this, helps us restrict the range for processes. The actual domain consists of fitness-relevant information and information that seems sufficiently similar to such information to make exposure to it trigger the system. The proper domain is a subset of the actual domain which contains only the fitness-relevant information (Sperber, 1996, 136, see also Buss, 2015, 45-46).

Here are some examples, based on things we saw in chapter 2. A display of two eyes gazing at the subject is in the proper domain of agency-detection. Two eyes painted on a wall are in the actual domain. Visual exposures to a twig and a stick insect that looks like a twig are in the actual domain of the agency-detector in birds, but only the display of the actual insect is in the proper domain. A face and a cloud that looks like a face trigger human facial-recognition and can induce an anthropomorphic inference. The face is in the proper and the cloud is in the actual domain. Finally, information about the proximity a caregiver has to a baby is in the proper domain of the attachment-system. Thoughts that indicate the level of psychological proximity a purported God has to a theist is in the actual domain of the attachment-system.

The last example illustrates that theistic concepts are within the actual domain of an evolved system and therefore within the range of information that bears on the process it implements. Given that these concepts are cultural, cultural information is in the actual domain of our systems. Dan Sperber (1996, see also Boyer, 2002) argues that human minds produce massive amounts of concepts, some of which increase their frequency as the result of our cognitive susceptibilities. As we have seen, minimally counterintuitive concepts (MCIs) of intentional agents have the best chances of achieving this since these are the most cognitively attractive (Ch. 2, sections 4 and 5). It is when such information becomes widespread that we

call it cultural. Then, in a feedback-process, cultural information "invades" the actual domain of our systems, with the result that mental epidemics arise (Sperber, 1996, 138-146). In light of the barrage of concepts our mind produces and which in turn might latch on to the functioning of numerous systems, the contours of the actual domains of these systems are unwieldy. Nevertheless, a distinction can be made between information that triggers a system and information that does not. A distinction can then be made between propositions the system must be reliable with respect to and propositions that can be discarded as irrelevant. Propositions tied to the informational content of the actual domain for the underlying system are relevant, and propositions tied to information outside that domain are not. In this way, the relevant range part of the generality-problem for processes can be solved. CSR contributes by helping us limit the range of information relevant to a given process.

Thus far, we have explored the theistic disposition (TD) from an angle that allows us to appreciate the distinct functional procedures it involves. CSR and Alston's psychological realism have helped us solve the relevant type problem and the relevant range problem for the processes specified above, thus providing a solution to the generality-problem for processes. Our finding that CSR helps us specify functional procedures and therefore also (given Alston's psychological realism) helps us identify a set of process-types by which theistic beliefs arise further undermines the claim that CSR does not explain how theistic beliefs arise. To be more specific, we now have grounds additional to those provided in chapter 3 to reject statements 3 and 4 in P3 in E-ENT, which contend that CSR does not explain by what processes beliefs form. Moreover, we have gained increased insight into how the TD looks from an angle that allows us to focus on distinct functional procedures.

However, given our argument for agent-reliabilism in the foregoing chapter, the crucial epistemic issue is not the reliability of processes. Reliability is necessary but not, as the True-temp cases indicate, sufficient for justification. Instead, justification requires that *epistemic virtue* is relevantly involved in the production of belief. Epistemic virtues are examples of *cognitive abilities*, that is, stable, reliable, integrated dispositions to produce true and to avoid false beliefs (Greco, 2012). As I have argued, the TD is a potential virtue. It is a gradually acquired, belief-forming habit which comprises the above-specified (and further) processes. We therefore now zoom out and consider the TD from an angle whereby we can better appreciate its nature as a complex belief-forming habit in its own right. It then becomes clear that the above processes, when occurring in subjects with the TD, can play a special role: they enhance and preserve the psychological relevance of the concept of God. Such occurrences are therefore what I called (Ch. 3) "relevance enhancing episodes". Such episodes lead to the formation of

the TD in the subject and they contribute to keep this habit going by preserving the relevance of God. A generality-challenge emerges once we turn our focus to the TD as a complex habit. What type of belief-forming disposition is this? Moreover, how can we restrict its range? The above solution to the generality-problem for processes is relevant to these issues, which we need to deal with before asking whether the TD is reliable the way virtue requires.

3. From the Theistic Disposition to the MCI disposition

The question we face in the present section is what type of belief-forming disposition the TD is. We need to know this to assess whether it is reliable the way virtue requires. We here encounter a generality-problem for belief-forming dispositions. This problem again consists of a two-pronged challenge: how to specify the right type and how to restrict its range appropriately. We continue to assume, as a central guiding idea, Alston's psychological realism. In particular, we employ the idea that we should seek a type which carves something psychologically distinct while leaving out irrelevant phenomena. What sort of psychology are we talking about? We saw above that by identifying functional procedures we found process-types. This, however, does not provide us with the type for the TD. This is because this disposition does not involve one functional procedure but many different procedures. Furthermore, the mere dynamics of these procedures do not fully capture what is distinct about the TD. The former involves proximate cognitive events whereby evolved structures make bets about fitness-relevant issues. The latter involves something else. When such betting-procedures occur in theistic subjects, they achieve something profound, namely that the intuitive concept of God becomes relevant. This is the result of how the concept entangles itself in the processing. It is the dynamics involved in these relevance-enhancing events we should seek, since it seems to be something profound, which is distinct from other psychological phenomena. In other words, the particular psychological phenomenon that the type must carve correctly is relevance-driven psychology. The type we seek covers belief-forming dispositions that involve this sort of psychology while leaving out dispositions that involve other forms of psychology. Let us now close in on the right type by envisaging disposition-types with different levels of generality and look for one that achieves this.

The first is "belief-forming disposition". This type covers literally all belief-forming dispositions no matter what processes they involve or what content the concepts have. The TD and all other belief-forming dispositions are tokens. Surely, this general type does not carve

something psychologically distinct while leaving out irrelevant phenomena. This type is therefore too broad.

Second, consider the type "belief-forming disposition which leads to beliefs about intentional agents". This type covers dispositions to form beliefs about any supposed supernatural intentional agent, including cognitively attractive and cognitively unattractive agents. The type therefore covers dispositions to form beliefs about intuitive versions of God, ghosts, spirits, witches, fairies, and demons. It also covers dispositions to form beliefs about highly counterintuitive agents. This includes theologically correct versions of God. Furthermore, the present type also covers dispositions to form beliefs about entirely mundane natural agents such as humans and animals. Very counterintuitive or entirely mundane concepts both fail to be memorable, catchy, and psychologically relevant the way relevance-driven psychology involves. Given that this type incorrectly covers such cases, it is also too broad.

We have now seen two broad types, which clearly are not the one we seek. Let us then instead start at the other end with narrow types and see whether that brings us closer to the right type. One could as a start propose super-narrow types that contain only one disposition. This means that each belief-forming disposition is a unique type. If we were looking for types that carved unique personality-structures, then we could perhaps envisage types with only one token. However, given that the type we seek carves dispositions involving relevance-driven psychology, this is problematic. Each personally unique type, which only has one token, fails to include countless dispositions that involve the same sort of psychology, a psychology whereby God becomes relevant in largely the same way. Let us instead consider moderately narrow types which cover distinct groups of tokens related to distinct versions of Abrahamic monotheism while leaving out everything else. Then, groups of dispositions to form beliefs about Yahweh, the Christian God, or Allah, each belong to distinct types. This provides us with types like "Jewish belief-forming disposition", "Christian belief-forming disposition", and "Muslim belief-forming disposition". We can envisage these tokens and their formation thus:

(a) S's exposure in context C to (c) (Allah) → a set of relevance enhancing processes involving (c) (Allah) → S's disposition to form beliefs about Allah.

(b) S*'s exposure in context C* to (c) (Yahweh) → a set of relevance enhancing processes involving (c) (Yahweh) → S*'s disposition to form beliefs about Yahweh.

This strategy is also problematic. Intuitive versions of the concepts of Allah and Yahweh are both cognitively attractive and can therefore both get a relevance-driven process started in the same way. There is no evidence for differences with respect to such psychology when it comes to culturally distinct versions of the concept of God. The type we seek must therefore include S's disposition and similar dispositions in C and S*'s disposition and similar dispositions in C*. Furthermore, it must include dispositions involving intuitive versions of the Christian concept of God. Let us therefore increase the scope of our type so that it involves all dispositions to form beliefs about God no matter which cultural variant of that concept they involve. We can then see the dispositions in (a) and (b) as culturally distinct tokens of the following type, which we have encountered many times:

(c) Exposure in a theistic cultural context to theistic concept → a set of relevance-enhancing processes involving theistic concept → Disposition to form theistic beliefs (TD).

We can consider the theistic disposition (TD) as a type in its own right. This type covers all dispositions to form beliefs about God. We here have a type that correctly includes culturally distinct dispositions that involve the God of Judaism, the Christian God, and Allah. Given that these all involve the same relevance-driven psychology, the current type seems initially promising. Is then the TD the right type? Notice that this type leaves out all dispositions that involve other deities than God. If this were correct, then all those dispositions would involve other sorts of psychology. But what is the evidence for this? Consider subjects in an ancient Aztec city in Mexico who are prone to form beliefs about the deities of the Aztec pantheon. Consider present day Hindu beliefs about a variety of deities. Are these propensities the outcome of a different sort of psychology? Does the type that includes all and only tokens of the TD carve a psychologically unique phenomenon? There is no evidence for this. There seems to be nothing psychologically special about God, the culturally distinct purported deity involved in Judaism, Christianity, and Islam. Rather, judging from the central contributions to CSR discussed in the present investigation, the evidence indicates that a wide range of cognitively attractive god-concepts can gradually embed themselves in intuitive thought, with the result that belief arises. This indicates that the sort of psychology involved when a subject in an ancient Aztec city becomes prone to form beliefs about The Feathered Serpent is the same as the sort of psychology involved when a subject in a Christian cultural context becomes prone

to form beliefs about Jesus.[102] The type we seek must therefore not fail to include the following disposition:

(d) S**'s exposure in context C** to (c) (the Feathered Serpent) → a set of relevance-enhancing processes involving (c) (the Feathered Serpent) → S**'s disposition to form beliefs about the Feathered Serpent.

Is the right type at least limited in scope to dispositions that lead to beliefs about cognitively attractive *deities* (gods)? Do we have grounds to consider something like "disposition to form beliefs about cognitively attractive deity" as the right type for the TD? This increase in scope brings us closer to the correct type given that it now is sufficiently wide to cover (d), and countless other dispositions unrelated to God. However, this increase in scope does not go far enough. It is a central contention in CSR that a wide range of *non-theistic* concepts of supernatural agents also are cognitively attractive and therefore can embed themselves in intuitive thought. This includes concepts of ghosts, spirits, demons, angels, witches, fairies, and ancestors (see Ch. 2, section 5 for details). Consider for example the account of the Fang-people that Boyer (2002) presents. Their beliefs arise as the result of the same relevance-driven psychology as the one involved in belief in deities (see Boyer, 2002, 58-105). Therefore, a type whose tokens include all (but only) beliefs about deities is, despite its massive scope, too narrow. The right type must also include the beliefs of the Fang and countless other beliefs in spirits, ghosts, demons, angels, and so on. It must for example include the following and similar tokens:

(e) S***'s Exposure in context C*** to (c) (the ancestors) → a set of relevance-enhancing processes involving (c) (the ancestors) → S***'s disposition to form beliefs about the ancestors.

[102] A caveat: in terms of relevance-driven psychology, these tokens belong to the same underlying type. If we instead seek types that carve psychological phenomena along other dimensions, then it could be that these two tokens belong to different types. For example, consider the ancient Aztec deity the "feathered Serpent" (Quetzalcoatl). It is represented in various artefacts as a rattlesnake with Feathers, hardly a trust-inducing image. Moreover, the worship of this and other Aztec deities, such as the Deity of Death, was associated with human sacrifice (see e.g. Austin and Lujân, 2008). In contrast, Jesus is in some Christian contexts, such as present-day evangelical communities, represented as a caring man who loves us just the way we are (Luhrmann, 2012). If we want to specify a type which for example covers all and only dispositions whereby the deity functions as what Kirkpatrick (2005) refers to as a "haven of safety", then, possibly, a right type should not include dispositions that lead to beliefs about the Feathered Serpent or the Deity of Death. But now that we look for a general type which covers all tokens of dispositions that involve *relevance-driven psychology*, a correct type must include both.

Given that S***'s disposition involves the same relevance-driven psychology as the tokens discussed earlier, a right type must include it. We now increase the scope so that we achieve this. What type do we obtain? This type covers beliefs about cognitively attractive supernatural agents in general, whether or not they are deities or supernatural agents of other kinds. This type covers dispositions to form beliefs involving the intuitive concept of God, the Feathered Serpent, Zeus, and countless concepts of spirits, demons, fairies, ancestors, and ghosts. These concepts all have the right structure which makes our minds susceptible to them. They are catchy and memorable but not highly complex and hard to grasp. Furthermore, they easily become activated in social interactions whereby evolved cognitive systems produce intuitions. In short, they are *minimally counter-intuitive* (MCI) concepts (see chapter 2, section 5). A type whose scope includes all and only dispositions to form beliefs about MCI agents seems to be the right one. This type carves a distinct psychological phenomenon, namely relevance-driven psychology. At the same time, it leaves out dispositions that arise and persist in other ways. An increase in scope beyond this would provide an incorrect type, given that we then would also include other psychological phenomena. Let us call this type the "MCI disposition". The dispositions in (a), (b), (d) and (e) are tokens of this type, which we can specify thus:

(f) Exposure in a given cultural context to MCI concept → a set of relevance-enhancing processes involving MCI concept → Disposition to form beliefs about purported MCI agent.

In conclusion, the right type for the theistic disposition (TD) is the MCI disposition.

We can now see how the TD is a culturally distinct token of the MCI disposition which arises in contexts where relevance-driven psychology yields belief in God rather than other deities. This relates to the need to introduce exposure to theistic ideas in order to provide an ontogeny for this habit (see Ch. 3, section 3). Notice that the MCI disposition in general then does not tend to produce beliefs about God. How representative are theistic beliefs of its output in general? It is hard to pin down the emergence of the MCI disposition as a distinct psychological phenomenon. The phenomenon that people form intuitive beliefs in spirits, ghosts, gods, and other purported cognitively attractive entities long predates the relatively recent emergence of Abrahamic monotheism in some parts of the world. It is probably quite ancient. For example, Matt Rossano (2009) dates religion in the form of shamanism and ancestor-worship to the Upper Paleolithic or even somewhat earlier. Moreover, many recent civilizations such as those of the Toltec, Olmec, Mayan, or Aztec people were not tied to Abrahamic culture and had other types of god-belief (see Baumard and Boyer, 2013). Consider

also Buddhist religious practice, in which belief in a range of MCI beings unrelated to God is widespread (see Pyysiäinen, 2009, 137-172 for examples). In light of this, it seems fair that a representative sample of beliefs produced by the MCI disposition covers cases that stem from various cultural sources stretching back some distance into our pre-Holocene past and from various present-day cultures, only a few of which are majority Jewish, Christian, or Muslim. This indicates that only a relatively small fraction of the beliefs we find in this sample are about the God of Abrahamic theism or other MCI beings tied culturally to belief in that specific deity.

Given that we now have identified the type to which the TD belongs, thus solving the relevant type problem, we can turn to the challenge of correctly limiting its range. This is crucial given our goal to establish whether this type is reliable the way virtue requires. To do this we need to specify a range of propositions it must be reliable with respect to.

Recall (Ch. 6, section 3), an epistemic virtue is a belief-forming disposition that tends to get things right and to avoid mistakes with respect to a range of relevant propositions when operating in a range of appropriate environments and conditions. The relevant range problem for belief-forming dispositions in general is thus the challenge of specifying these ranges correctly. The present challenge is doing this with respect to the MCI disposition. What is an appropriate range of environments, conditions, and propositions for this disposition?

When it comes to the first, a similar restriction to the one for process-types (discussed in section 2) seems correct. That means that we keep out evil demon scenarios. Appropriate scenarios obtain in close possible worlds but evil demon scenarios (presumably) obtain only in remote worlds. Furthermore, we should allow a broad range of conditions to apply except those that obviously are inappropriate for the disposition. What is an example of such a condition? John Greco describes some conditions that are obviously inappropriate for a given ability. For example, the condition of being drugged or injured when one is about to perform as a baseball player is inappropriate. Therefore, failure to perform well in such a condition does not undermine the baseball player's ability (Greco, 2007, 213-215). What would be analogous to this with respect to the MCI disposition? A condition in which a subject has brain damage to those areas that govern Theory of Mind related tasks seems inappropriate. Inability to form intuitive beliefs about the mental states of MCI agents disables the MCI disposition in an inappropriate manner. For example, consider a subject who in the actual world has the MCI disposition but who in a close world is brain-damaged and therefore fails to believe in an MCI agent who actually exists. This failure to get things right should not affect our assessment of the reliability of his disposition. This is because such a condition is inappropriate.

What environments and conditions are in contrast appropriate? Recall how culturally contingent the MCI disposition is. We find different tokens of it in different cultural contexts where minds that operate in the same way respond to socially transmitted pieces of cognitively attractive information. The variety of cultural contexts in which such information is transmitted and relevance-driven psychology therefore occurs is enormous (see e.g. Sperber, 1996). This indicates that within an appropriate range of environments, we must situate a very large number of social and cultural contexts where subjects are exposed to countless variants of cognitively attractive information. When it comes to conditions, we can also here allow a very wide range. The crucial issue is how the MCI disposition tends to operate in conditions in which the mind *works normally* during many functionally different kinds of social interactions with conspecifics as well as during encounters with predators and prey.

Finally, we need to consider the set of propositions with respect to which the disposition must be reliable. For different candidate (epistemic) abilities, there are different sets of propositions with respect to which they should be (globally) reliable. If p is in a wider set of propositions relevant for candidate ability X, then S must not only get p right but tend, in a range of close possible worlds, to get propositions in this wider set right (see Ch. 6, section 3). This means that the reliability we will assess is global: it pertains to a wider range of propositions rather than just one. Notice also that it is not sufficient for reliability that one gets things right in the actual world. What matters is average performance in a range of appropriate environments and conditions which obtain in a set of close possible worlds. The scope of the MCI disposition gives us the relevant set of propositions. We found this type by focusing on just how far-reaching relevance-driven psychology is. As we have seen, MCI concepts get such a psychology started. When such concepts increase their relevance in minds, deities (such as God or Zeus) and non-deities (like ghosts or witches) begin to seem real to subjects. Subjects then are prone to form beliefs about such entities. Given that the MCI disposition is a disposition to form beliefs about MCI beings rather than other things, it has a domain, namely that of purported MCI beings. The set of propositions it must get right pertains to the contents of that domain. Just as the type we specified includes all and only dispositions that involve relevance-driven psychology, the set relevant to this type includes all and only propositions about MCI agents. Given the massive scope of the MCI disposition, which leads to so many different beliefs in different contexts, this set is massively large. Just imagine the endless number of culturally distinct variants of MCI concepts. Moreover, the propositions in this set are not only about the existence of MCI agents. Recall our discussion of the primacy of practical, situational religious beliefs (Ch. 3, section 3). It indicates that the output of the MCI disposition consists

of beliefs about the presence or absence of MCI-agents in concrete situations, and what intentions, goals, or purposes they have in relation to the subjects. Within the relevant domain of propositions for this disposition we can therefore place propositions about these practical, contingent issues. This then is the set of propositions the MCI disposition must tend to get right in a range of relevant possible worlds. We now ask whether this type achieves this and therefore is a virtue or not.

4. Troubled navigation: the argument against epistemic virtue

The massively large set of propositions relevant to the MCI disposition is a landscape in which this disposition navigates efficiently if reliable. A reliable belief-forming disposition navigates this landscape by tending (in a range of appropriate environments) to form beliefs in those propositions that are true (if any of them are) and to abstain from forming belief in those that are false. If the truth-value of any proposition changes, then the disposition responds appropriately. If no propositions in the MCI landscape are true, then it abstains from producing any belief. It is then on standby but *would* produce belief if an MCI agent came into existence. Is the MCI disposition reliable in this way? Does it meet this necessary requirement for being an epistemic virtue? We can plausibly deny this, even on assumptions theists agree with. Here is why.

Consider the set of propositions the MCI disposition must get largely right. Must we be agnostic about the likely truth-value of its contents? This seems wrong. It borders on the bizarre to suppose that *most* MCI beings are real. Just consider what reality we would live in if that were true. Thousands of deities exist, our world is populated by massive amounts of benign and hostile intentional creatures that swirl around us, haunt our houses, and lurk in the forests, there are fairies, trolls, witches, elves and demons, and many other supernatural things. We do not live in that reality. But this means that most of the propositions in the above-specified set are false. Then, since the MCI disposition nevertheless fails to abstain from forming beliefs in MCI agents, it is not reliable the way virtue requires:

The argument against epistemic virtue

P1. It is necessary for the virtue of the MCI disposition that it tends to produce true and to avoid false beliefs with respect to the set of MCI propositions in a range of appropriate cultural environments and conditions.[103]

P2. When occurring in a wide range of appropriate cultural environments and conditions, the MCI disposition widely produces beliefs about MCI agents.

P3. Most MCI agents do not exist.

C. The MCI disposition fails to meet a necessary requirement for being an epistemic virtue.

Notice that P3 does not presuppose or logically entail C, since (as mentioned above) the MCI disposition could achieve reliability in virtue of *abstaining* from forming false belief. This means the MCI disposition could be reliable in the right way even if P3 were true. The combination of P3 with a premise P2*, stating that the MCI disposition widely *abstains* from producing beliefs about MCI agents but would embark on MCI production if MCI beings came into existence, could even be seen as supporting the *denial* of C. The reason is that a disposition with this property seems reliable the way virtue requires. However, there is no evidence that the MCI disposition functions that way. By contrast, in the argument above the specific combination of P3 with *P2* tells us that the disposition fails to be reliable in the right way. Given that this reliability is necessary for virtue (P1), it follows that the MCI disposition is not a virtue.

Notice also that P3 is true even if *some* MCI beings exist. This is because the denial of the claim that most MCI beings exist is consistent with the thesis that some of them do exist. Theists in fact widely agree with this denial given that they are prone to form beliefs about *God* but not most MCI beings. Those theists who ponder this reflectively will also agree: most MCI beings surely do not exist. God exists.[104] Consider for example the Islamic creed that there is *no other deity than God*, the Jewish confession of faith according to which *God is one*, and the

[103] This is the reliability-requirement for virtue defended by Greco (2007). Recall, epistemic virtues are cognitive abilities, and such abilities involve a tendency obtain true and to avoid false beliefs with respect to a correctly specified set of propositions, in a range of relevant close worlds (see e.g. Greco, 2007, 206-216). For a defence of this virtue-epistemological model, see Ch. 6, section 3 and 4.

[104] The reflective theist might also go further and claim that *no* MCI agent exists and point out that God, correctly understood, is no such agent. But this does not undermine the arguments above. Suppose God is no MCI agent and that God exists. It is then still true that most MCI agents do not exist. But then, the MCI disposition is not a virtue, and the theistic beliefs it produces in theistic cultural contexts are still unjustified. That leaves open the possibility that some reflective theists have formed theistic beliefs in other and more refined ways, through other processes than those we have analysed. This also does not challenge the argument, since the processes CSR has revealed are not necessary conditions for the production of theistic belief (for this, see the clarifications in Ch. 1, section 5, and the discussion of "personal explanation" in Ch. 3, section 2). I did not argue that it is in principle impossible for theists to form beliefs about God by basing such beliefs on carefully reasoned grounds or carrying out other intellectual procedures. I argued that in typical cases, theistic beliefs arise through the processes CSR reveals. These are the processes the MCI disposition involves, which lead to distinct religious beliefs in distinct cultural contexts.

Christian doctrine that God is *one divine substance* consisting of Father, Son, and Holy Spirit. The claim that there are no other deities than God and that polytheism therefore is false is central to these theistic doctrines. But even those theists who are open about the existence of a range of supernatural agents besides God (for instance, angels and demons) probably do not believe that *most* such agents, independently of cultural context, exist. One need only imagine the massive amount of such concepts that human minds have produced through history to see this. So, theists need not and will probably not reject P3.

Given C we can now propose an argument against the epistemic justification of three types of theistic beliefs:

The argument against epistemic justification

P1. It is necessary for the epistemic justification[105] of beliefs that an epistemic virtue has been relevantly involved in their formation.[106]

P2. In theistic cultural contexts, the MCI disposition leads to intuitive, explanatory, and folk-theological beliefs about God.[107]

P3. The MCI disposition fails to meet a necessary requirement for being an epistemic virtue.

C. The intuitive, explanatory, and folk-theological beliefs the MCI disposition leads to are epistemically unjustified.

We will now see that even if God or a limited set of MCI agents do in fact exist, the above verdicts are still correct.

Suppose first that there are some lesser deities[108], ghosts, and spirits, but that God does not exist. Those who in the actual world believe in the three former entities get things right but fail to abstain from false belief in many close possible worlds. There are two ways in which this can happen. The first is that in many close possible worlds, the three former entities are *not* real, but the subjects, as the result of the same relevance-driven psychology in the same cultural environment, believe in them nevertheless. The second way is that in at least some close possible worlds the subjects inhabit theistic cultural contexts and therefore, as result of the same psychological processes, believe in God who (we now suppose) does not exist. In both cases,

[105] I.e. the doxastic justification. See chapter 4, section 3 for clarification.
[106] For a defence, see chapter 6, sections 3 and 4.
[107] For a defence, see chapter 2, chapter 3, section 5, and section 3 of the present chapter.
[108] By a lesser deity, I mean a deity who is not omnipotent or logically necessary (like God is said to be). It seems that most purported deities are lesser.

failure to abstain from false beliefs in close possible worlds shows that the MCI disposition is not reliable, even if some MCI beings exist.

Suppose instead, as theists believe, that God and a few culturally related beings (like angels or saints perhaps) exist. Now, those who in the actual world believe in God (and angels and saints) get things right by believing in the fraction of supernatural agents that exist. Two factors nevertheless indicate their lack of reliability. The first is that they (in the actual world) get crucial properties of God (like omniscience and omnipotence) wrong by anthropomorphizing him.[109] The other is that in many close possible worlds, the subjects inhabit cultural environments where belief in non-existing deities or spirits is common, and here the same relevance-driven psychology leads them to believe in those entities instead of God, who (we now suppose) exists. Their failure to abstain from false belief in the actual and in many close possible worlds again shows that the MCI disposition is not reliable, even if some MCI beings exist.

The problem in these two cases is not that in a possible world where *God* does not exist the subjects would nevertheless believe in God.[110] I am not arguing that theistic beliefs are *insensitive.* S's belief that p is sensitive if and only if the following two statements are true: (1) if p were false, then S would not believe p, and (2) if p were true, then S would believe p. This means for example that S's belief that p is not sensitive if in the closest possible world in which p is false, S believes that p (Prichard, 2010, 732-733). It seems (as the first case shows) that beliefs in lesser deities and spirits are insensitive. It is not true that if such entities existed, then subjects would believe in them. Rather, subjects would believe in them if relevance-driven psychology led to belief. It is also false that if these entities did not exist, then subjects would not believe in them. The decisive factor that accounts for belief or lack thereof is whether the concepts of such entities become relevant, not whether they exist. This follows from my argument in earlier parts of the investigation. As I have argued, we can explain in a completely naturalistic fashion how these kinds of beliefs arise.[111] In contrast, it is not entirely clear that belief in God is insensitive in this specific manner, even though we can, as I have argued, explain also such belief naturalistically. The reason has to do with a certain modal property God is said to have, namely the property of being a logically necessary being. This creates certain

[109] For my argument for this, see chapter 5, section 3.
[110] Murray (2010, 175) and Clark and Rabinowith (2018, 118) discuss arguments that involve this claim.
[111] Chapters 2 and 3 argue that CSR can explain how beliefs in God arise. Given the argument in the present chapter that the theistic disposition is a token of a more general disposition which leads to a wide range of MCI beliefs, we can also now suppose that CSR can explain how such beliefs arise. In fact, the explanation of how theistic beliefs arise seems to be just a token of a broader explanation of how MCI beliefs arise.

complications, because it means that if God exists, then he exists in all logically possible worlds. Then, it is not clear how the antecedent in (1) above could be true[112] and how we coherently could state that there is a closest world in which God does not exist but where the subject believes in God. This complication, however, does not affect the argument that the MCI disposition fails to be (globally) reliable, which is the decisive issue.

In conclusion, the MCI disposition fails to be reliable the way virtue requires and therefore, the theistic beliefs it leads to lack epistemic justification, even if God exists.

We shall now consider two possible responses to the arguments above.

The first concerns a respect in which theists, if God exists, could not easily be wrong, despite the truth of what I have argued. One might suspect that the fact that they could not easily be wrong in this respect undermines my verdict. Here is the sense in which theists could not easily be wrong. If God exists in the actual world, then, given his logical necessity, God exists in all close possible worlds. This means that in all close worlds in which a subject S believes in God, S's belief that God exists is true. S then could not easily be mistaken about God in worlds where S believes in God. It seems that S's belief in God therefore is *safe*, given the following specification of safety, provided by Prichard (2010, 734). S has a safe belief that p if, and only if, in most nearby possible worlds in which S believes that p, p is true. This means that S could not, when believing that p, easily be wrong. If God exists, then this seems true of theistic beliefs. In contrast, this does not seem true of S's true belief that a ghost haunts his basement. This is because in many close possible worlds, S believes that this ghost haunts his basement, but it actually haunts other basements. Therefore, whilst theistic beliefs could not easily be false (if theism is true), beliefs about ghosts and many other MCI agents could easily be false (even if true in the actual world). It then seems relevant to ask whether this difference between the modal situation for theistic and other religious beliefs indicates that the negative verdict above concerning theistic beliefs is wrong.

In response, let us start by noticing that the safety (in the above-specified sense) of theistic beliefs (if God exists) is consistent with the premise that a *globally* unreliable and therefore unvirtuous disposition is responsible for their production. It is possible for a globally unreliable disposition to output some true beliefs that are safe. The issue is whether their safety challenges the verdict that they are unjustified. Then, a good argument should be able to show that it is not necessary (as P1 in the argument against virtue states) for the justification of beliefs that an epistemic virtue is involved in their formation. In other words, this turns into a discussion

[112] As Goldman (1986, 48) points out, if p is a logical necessity, then the antecedent in the first subjunctive conditional above is logically impossible.

about the correctness of agent-reliabilism. Given that we have already defended this view, we shall not return to that discussion here. Instead, here is another and good reason to deny the claim that the safety of a belief suffices for its justification. Consider the following points which Ernest Sosa makes about beliefs in necessary truths. Sosa starts by pointing out that if one forms the belief that a necessarily true proposition p is true, then one could not easily be wrong when forming that belief. This is because that proposition is true in all logically possible worlds. Therefore, in all worlds in which one believes that p, one is right about p. However, this is not sufficient for the justification of this belief, Sosa argues next. We can imagine all kinds of problematic ways in which subjects form beliefs in necessarily true propositions. Imagine for example that they formed such beliefs based on wishful thinking. This is not sufficient for justification (Sosa, in Conee and Feldman, 2011, 208-210, cf. Sosa, 2007, 29). This seems correct. This means the safety of belief in God - if God exists - fails to provide grounds to deny the negative verdicts above. Safe belief in God formed in an unvirtuous way, i.e. formed by the MCI disposition, is not epistemically justified.

Here is another possible objection to the argument against virtue, based on a remark Alvin Plantinga (2000) makes about the contingency of many of our beliefs. Considering it allows me to clarify something important about what I argue and what I do not argue. Plantinga points out that we could easily have failed to believe many of the things we do believe in religion but also in science and philosophy. That is, in a range of close possible worlds where we live in other cultural and social contexts, we fail to have the beliefs we presently have about a variety of matters. For example, Plantinga suggests, if Einstein had been born in another century, he would probably not have believed in the theory of special relativity. Einstein therefore could have easily failed to believe in that theory. Another example is that a subject who is a Christian in the actual world is a Muslim in a close possible world in which he grows up in an Islamic culture. This does not show us that these beliefs are unjustified, Plantinga suggests. The reason is that the following principle, which leads to this negative verdict, is wrong: if S's belief that p is such that if S had lived in another context, then S would not have believed that p, then S's belief that p is not justified. Plantinga's objection to this principle is that it is self-defeating. In many close possible worlds, we fail to believe in it. Then, by this principle, we should not accept this principle (Plantinga, 2000, 427-429).

Here is the objection. P2 in the argument against virtue states that the MCI disposition widely leads to different beliefs in different contexts. This means it is a culturally contingent disposition. Then, one might think, the argument appeals to something like Plantinga's principle. Because this principle is problematic, the argument against virtue is problematic.

In response, Plantinga is right that many of our beliefs are contingent. For example, those who are Christians in the actual world are probably Muslims in close worlds. Consider also this. A subject who in the actual world has a reasoned belief in evolution probably fails to have this belief in many close possible worlds where he or she has been exposed to creationistic ideas and failed to obtain an education. Given the above principle, then the Christian beliefs and the reasoned evolution beliefs are both unjustified simply because the subjects could easily have failed to have them. This is problematic, but not what I have argued.

I argued that justification depends on whether virtue is relevantly involved in the formation of belief. The virtue of a given belief-forming disposition hinges on its average performance across relevant worlds. A measure of the performance of the disposition therefore does not hinge on what happens in worlds where it is not formed. Facts about what one would believe as the result of whatever disposition therefore does not tell us about the epistemic status of our current beliefs. Suppose S through a deliberately acquired belief-forming disposition D forms a reasoned belief in evolution. Furthermore, suppose that in many close worlds, S fails to acquire D and instead, as the result of exposure to creationistic thinking and relevance-driven psychology, becomes a creationist. That the possibility of creationist belief looms for S does not tell us that D, which produced belief in evolution in the actual world, is unreliable. Instead, it tells us that D is a precarious disposition which requires learning and careful thought. To be able to form a reasoned evolution belief is a privilege we cannot take for granted.[113]

In contrast, the situation for the Christian who could easily have been a Muslim does tell us something epistemically negative. This is because we here have grounds to believe that the same underlying (MCI) disposition leads to these different results in a range of worlds. As we have seen, among the results we should envisage, we can include beliefs in Zeus, ghosts, demons, angels, or witches. What creates the epistemic problem is not the mere contingency of the beliefs that arise but the unreliability of the responsible disposition.

In conclusion, these two responses fail, and the conclusion remains: The MCI disposition is not a virtue, and the theistic beliefs it leads to therefore lack justification.

Conclusion

[113] As Kelemen (2004, 299) argues, "promiscuous teleology", a tendency to over-attribute design and purpose to the natural world, makes children disposed to creationist thought. In a discussion of implications for science-education, Kelemen further suggests that the teleology-effect contributes to explaining why children often fail to understand evolution. This supports the contention about how precarious the formation of scientific beliefs is.

We have thoroughly investigated the theistic disposition by viewing it from different angles, including one where we focused on proximate events related to functional procedures, and one where we focused on what general type of disposition it is. The TD, which we found to be a token of a broader belief-forming disposition, could have been but, to judge by the best available evidence in CSR, is not an epistemic virtue. If a subject has this disposition, then he or she does not tend to get things right and to avoid mistakes in the actual and in close possible worlds when in appropriate conditions. The theistic beliefs that arise from the theistic disposition are therefore not justified. This is because this disposition fails to meet necessary requirements for being an epistemic virtue, and, as we have seen earlier, the justification of beliefs is derivative on the epistemic status of the dispositions from which they arise.

Given how extensively CSR contributes to showing this, such research has normative implications for what I call theist religion in sense (3). Recall that relevance in an epistemic evaluation of theistic beliefs suffices for such implications. The implications we have established are relatively strong: the evaluation led to the conclusion that *theistic beliefs are not justified*. We have thus moved beyond our earlier conclusion that two central arguments for ENT (the epistemic neutrality thesis) fail and now concluded that CSR has negative implications for the epistemic status of theistic beliefs. Of the three main positions we have considered, namely neutrality, positive relevance, and debunking, the latter is therefore correct.

This concludes the present chapter and the third part of the investigation. The final chapter takes stock of what we have seen in the present investigation as a whole.

Chapter 8. The implications Cognitive Science of Religion has for Theist Religion

1. Introduction and chapter-outline

Empirical science helps bring about philosophical progress by providing starting-points for theorizing and by contributing to answering normative questions. We should philosophize in a landscape science provides but without thereby leaving philosophy behind. But where do we, if that is even possible or desirable, draw a line that separates the two? A philosophy of religion strongly informed by cognitive science must navigate in this blurry landscape. For those engaged in empirical scholarship of religion, the outcome might seem too philosophical, but for those engaged in philosophy, it could come across as too focused on the empirical. I believe we can diagnose these impressions as the result of too much belief in a neat division of intellectual labour in the study of religion. Empirical scholarship is for example often said to bracket philosophical issues while philosophers are depicted as armchair scholars who work out the answers to abstract questions of little empirical relevance. This picture is not accurate. Empirical scholarship contains philosophy, albeit often in a non-explicit form. Philosophers do not, or at least should not, cut off what they do from empirical scholarship. The present investigation has - from a theoretical perspective where empirical findings in cognitive science play a prominent role - revealed important implications of the cognitive science of religion. We are now ready to take stock of these.

I distinguished three types of normative implications for what I call theist religion. CSR has implications of the first type if the truth of its central claims bears logically on the truth of theism. One way in which such implications obtain is by logical incompatibility. The science and religion scholar might therefore want to scrutinize whether CSR and theism are compatible or at least can be re-interpreted in such a fashion. This is an interesting and important topic. The present investigation nevertheless chose another focus and presupposed that central contributions to CSR are or at least can be logically compatible with theism. This presupposition (which is not uncontroversial or obviously true) gave us space to focus extensively on two further types of implications, both of which deserve independent attention.

The first of these concerns the relevance CSR has in assessment of theistic arguments and theses. The second concerns the relevance such research has in evaluation of the epistemic status of theistic beliefs. What implications did the investigation reveal? Section 2 of the present chapter provides a summary of the three main parts of the present investigation and brings out the most important findings and conclusions. We then in the third and final section return to the notion of the "unweaving of religion" (see Ch. 1, section 1), which is based on Richard Dawkins' notion of the unweaving of the rainbow. Richard Dawkins and Daniel Dennett agree that we both can and should unweave religion, or in Dennett's words, "break the spell". Furthermore, CSR-scholarship figures as part of their criticism of religion. Focusing on Dawkins, we ask what sort of role such scholarship plays in this criticism. Contrary to what has recently been suggested, this criticism does not involve an argument from CSR to atheism. It instead involves an argumentative strategy whereby such research is used as a helpful *auxiliary* tool. We can add this to the stock of ways in which CSR can be of use in the philosophy of religion, in particular by providing an aid for the critic of religion.

2. Taking stock of implications for theistic theses and beliefs

Where do concepts of God, Zeus, spirits, ghosts, ancestors, and so on, ultimately come from? Why are we susceptible to these concepts, and by what process do some of them become part of resilient belief-structures that guide how subjects live and even how whole societies are run? With these kinds of questions, we started part 1 of our investigation of the cognitive science approach to religion. Our best guess based on my selection of influential work in this field is that the ultimate origin of religious concepts is found in the cognitive processing of ancestral minds dealing with adaptive challenges like detecting predators and prey and handling social interaction with conspecifics. False positives, the high frequency of which is a cost evolution allows since it enables the organism to avoid costlier false negatives, gave rise to ideas suggesting that there might be invisible agents around. A cultural selection process, which the biases in our minds influence, have then gradually increased the frequency of particular variants of such thoughts. These variants are successful because of how well they exploit our cognitive susceptibilities. In particular, we are especially susceptible to minimally counterintuitive (MCI) concepts, which are memorable and readily embed themselves in intuitive thinking. The intuitive concept of God is a relatively recent and still influential example of such concepts.

Concepts of the deities in the Aztec pantheon and the many concepts of spirits, ghosts, and ancestors are other examples.

Humans are cultural learners who crucially depend on information provided by others in the social context. We have seen how biases nudge us to pay special attention to the behaviour of others in our particular context. We are conformists who tend to follow the crowd and who are inclined to imitate the successful. This contributes to explaining why people tend to believe in the specific deities that are central to their social and cultural context. The theistic disposition (TD) arises in social and cultural contexts where subjects are exposed to credibility-enhancing displays of belief in God. The formation of the TD firstly requires that the mind is exposed to socially transmitted information on the basis of which the mind (usually of a child) constructs that particular representation. This representation consists partly of naïve ontological assumptions and partly also of socially transmitted information. The formation of the TD secondly requires that this representation gradually becomes more and more relevant to the subject. This makes the subject increasingly prone to believe in God. By providing this account, chapters 2 and 3 defended the thesis that CSR, a loosely delineated area in the study of religion which combines cognitive psychology, evolutionary theory, and anthropology in an attempt to explain religion as a natural phenomenon, *can* explain how people become theists and how theistic beliefs arise.[114] We called this the explanatory adequacy thesis:

> CSR can explain how people become theists and how theistic beliefs arise.

We saw how this research combines ultimate and *proximate* perspectives and thus provides important insights about the "nuts and bolts" of religion, not only an ultimate account in terms of cultural evolution. It follows from the account of the TD that to be a theist is not just to believe that God exists, but it is to be prone to think of God in a large variety of contexts. Moreover, such a propensity has the effect of making God seem real. The same is true of similar belief-forming propensities in other contexts. The minds of those who are prone to form beliefs about Zeus bring about the seeming reality of Zeus. The minds of the Fang-people that Boyer describes bring about the seeming reality of the ancestor-spirits. Finally, the minds of those who are prone to think of God make God seem real. I argued that by accounting for the general process by which subjects become theists, CSR also by extension contributes to accounting for individual cases, which can be seen as tokens of that process. This indicates that cognitive

[114] An important assumption is that central contributions to such research are *true or at least highly plausible.*

science rather than religious confabulatory narratives provides our best shot (thus far) at an explanation of how individual subjects become religious. Given that we can in a naturalistic fashion explain the phenomenon that people believe in deities, the existence of such entities is explanatorily irrelevant. The truth of the explanatory adequacy thesis also shows us something else: the central premise of the externalist argument for the epistemic neutrality thesis is wrong. That premise states that CSR cannot explain how people become theists or how theistic beliefs arise. For a summary and discussion, I refer the reader to the concluding remarks in chapter 3 and to chapter 6, section 1. The central point is that the explanatory adequacy thesis strongly undermines the externalist argument for the thesis that CSR is epistemically neutral.

Part 2 of the investigation explored the relevance CSR has in assessment of philosophical arguments connected to theism. Such relevance establishes normative implications in sense (2) for what I call theist religion. The two main philosophical positions we considered was theistic evidentialism (TE) and reformed epistemology (RE). Both are theses connected to theism, and they also contain particular arguments for the existence of God or for the warrant of theistic beliefs. TE is a conjunction of two claims. The first is an internalist thesis according to which only mental internal states have epistemic relevance. CSR does not undermine this specific view, but good arguments in the current epistemological literature indicate strongly that another thesis is correct, namely externalism.

The second claim in TE is the thesis that belief in God is justified by good evidence. CSR undermines this claim to some extent by having negative relevance for at least two moral arguments for the existence of God. As John Teehan has shown, such research also creates problems for theistic accommodation-theses according to which CSR has discovered how God sculpted our cognition to enable belief in him. Since we assumed that CSR and theism are or at least can be logically compatible, we did not go into the issue of a possible logical conflict here. I argued that what the cognitive problem of evil at least does, given that assumption, is to encumber theist philosophy with an added explanatory burden. When such a burden increases in the light of CSR, such scholarship makes theistic accommodation of CSR harder. On this interpretation, and given the bearing on the moral arguments, we found normative implications: CSR negatively affects the balance of the evidence on which the justification of belief in God hinges according to TE. Such research undermines but does not decisively rebut claim 2 in TE.

We found in chapter 5 that CSR has serious impact on the thesis that we have a god-faculty. Given that this is a central part of reformed epistemology (RE), such research therefore undermines this thesis. (For a more extensive summary, see the concluding remarks in chapter 5.) The central claim was that advocacy of the god-faculty thesis faces a dilemma in the need

to choose between two main interpretations of what this faculty is. On both interpretations, the thesis becomes problematic. On the first, there is a clash with CSR. On the second, one needs to depart from a central tenet of RE, the thesis that we can know God *without* having to reflect or reason. We thus found CSR to have serious impact on another theistic thesis, and this provides another example of normative implications in sense (2) for theist religion.

The main conclusion based on what we have seen in part 2 of the investigation is the following:

> CSR has strongly negative but only moderately extensive implications for theistic arguments and theses.

By this, I mean that such research has strong impact on specific arguments but not a far-reaching relevance for theistic arguments and theses in general (for more, see the concluding remarks in chapter 5).

We now turn to implications in sense (3). Given the problems *neutrality* and *positive relevance* face, we saw that debunking turned out to be the most viable option. Our evaluation of the theistic disposition in chapters 6 and 7 then bore the truth of that option out. We found that the TD is a culturally specific token of a broader belief forming disposition, the MCI disposition, which leads to different religious beliefs in different contexts. Our cognition makes us susceptible to form a wide range of religious beliefs, but which specific beliefs subjects end up as prone to believe in, depends on the cultural environment they happen to inhabit. Belief in a given MCI agent arises if the concept of that agent increases its relevance in the mind of the subject. We can account for how this happens without presupposing that any MCI agent exists. This indicates that the MCI disposition does not, to use the term Griffiths and Wilkins (2012) employ, *track religious truth.* Instead of helping subjects navigate in the variegated landscape of religious belief, this disposition leads subjects to form belief in the particular deity of their context, independently of whether any proposition in the MCI set is true. The central premise in the *argument against virtue* was nevertheless slightly different. Both theists and non-theists agree that most of the propositions relevant to the MCI disposition cannot possibly be true. An epistemic problem therefore arises from the manifest and widespread failure of subjects to abstain from belief in them. This failure shows that this type of disposition does not tend to get things right about the propositions in the MCI set in the actual world, and in a range of relevant possible worlds where subjects inhabit a wide range of social and cultural environments. In other words, this disposition fails to be reliable the way virtue requires. *The argument against*

justification showed that the theistic beliefs the MCI disposition leads to therefore lack justification.

Input from CSR figured into both the preparation for these arguments and provided insights that motivated central premises. For example, the notion of the theistic disposition (chapter 3), our scrutiny of it from the angle of functional procedures and from the angle of disposition-types (chapter 7), the argument that it is a token of the MCI disposition (chapter 7, section 3), and the argument that this disposition is strongly culturally contingent (chapter 2, sections 5 and 6, chapter 7, section 3) are all strongly informed by current research in CSR. We can conclude, given the normative evaluative framework presented in chapter 6, that CSR contributes strongly to show that theistic beliefs in representative cases lack epistemic justification. On this basis, our main conclusion in part 3 of the investigation is that,

> CSR has strongly negative implications for the epistemic status of theistic beliefs.

This means that CSR has strong implications for what I call theist religion in sense (3).

How strongly do these normative implications bear on the prospect of justified theistic belief in general? Notice that the implications we found do not primarily reflect the impact cognitive science has on theistic arguments. Then, beliefs about God based on arguments, if there are any, are not strongly affected. I did not try to show that CSR conclusively nudges the balance of the evidence in disfavour of theism. Furthermore, I did not argue that it is not possible for theists to base beliefs on arguments. Instead, what I claimed was that theists in representative cases do not base their theistic beliefs on arguments, and personal narratives about the formation of religious beliefs are in many cases confabulations (see Ch. 3, section 2). It does not follow that to base theistic beliefs on arguments is *impossible*. It is therefore an open question whether at least some theistic beliefs are justified by being based on arguments, and thus presumably arise in a way that the account of the MCI disposition does not capture. We should nevertheless appreciate that the conclusion above is a substantive view on the implications of CSR, which strongly departs from what seems to be the current consensus: that something close to neutrality is correct.

Internalists will of course object to the normative framework involved in the argument. They might for example contend that lack of reliability is compatible with justification, and that unless we grant externalism, the conclusions do not follow. This seems right. The arguments are externalist and do assume that justification requires the relevant involvement of a reliable cognitive virtue. But it is still, I would think, a substantive view that *given externalism,* which

after all is a mainstream thesis in current epistemology, such negative conclusions about the epistemic status of theistic beliefs are correct.

Are the verdicts too harsh given how broadly they apply? For example, they apply to subjects in contexts where information about cognition and reliability is unavailable. Consider a subject in a medieval Christian town who, like everybody else, believes in God. This subject obtains the concept of God at an early age and there is frequent exposure to credibility-enhancing displays of belief in God both by parents and socially influential leaders. Is it too harsh to say that the religious beliefs of this subject lack justification? This subject does nothing obviously objectionable when forming this belief. In response, we should recognize that while there are some epistemically positive aspects of the beliefs this subject forms, justification is not one of them. As Goldman (1988) argues, we should distinguish justified belief and merely blameless belief. The former is a well-formed (reliably formed) belief. The latter is an "ill-formed" (unreliably formed) belief that we cannot plausibly blame the subject for forming. As an example of the latter, Goldman mentions a subject S who arrives at a belief through astrology, a well-established practice in S's culture and whose unreliability S is unaware of (Goldman, 1988, 52-60). John Greco also argues that there can be something epistemically positive about beliefs formed in an unreliable fashion. This is the case if those beliefs seem entirely appropriate to the subject, who has no way of finding out that they are false or that they arise from an unreliable process. This is nevertheless not sufficient for "positive epistemic status", i.e. something like doxastic justification, warrant, or well-foundedness (Greco, 1993, 416-419). In light of this, it seems possible that theists are, at least in some contexts where there are no indications that their beliefs arise from an unreliable disposition, blameless in forming beliefs about God. We can recognize this without departing from the verdict that their beliefs are not justified.

3. Armchair criticism of religion and how philosophical naturalists can go further

In introducing the current investigation, I pointed to a sense of uneasiness many feel about the prospect of a naturalistic science of religion, for example among those who themselves have religious beliefs but even more strongly among those who outright reject the idea that we can explain religion. Those who insist that the only legitimate way to pursue the unwieldy phenomenon of religion is through hermeneutical methods face an irreconcilable conflict with

CSR. We initially asked whether the theist faces a conflict of similar seriousness, a situation where a choice between remaining a theist or accepting central theories in CSR *must* be made. We then allowed as an assumption that CSR and theism are or at least can be seen as logically compatible. This means we granted, for the purposes of the present investigation, the falsity of a central premise in the conflict-model (Ch. 1, section 3), namely the notion that science and religion are irreconcilable. But the uneasiness persisted despite this. We found that even if we concede compatibility, CSR and what I call theist religion have a disharmonious relationship: they do not go well together, and it is hard to make them agree.

This final section considers an *additional* way in which one can appeal to CSR in criticism of religion. Here CSR plays an auxiliary role in advocacy for atheism. We shall see that this is a legitimate strategy to pursue for the critic of religion, which is independent of those the current investigation has developed and pursued. This brings us back to where we started our investigation. Richard Dawkins (1998) argues that science in general does not belittle what it explains. Instead, by providing insights about natural phenomena, science inspires awe and poetry. Dawkins seems to find an anomaly to this general pattern in the case of religion, given his appeal to CSR in his (2007) attack on religion. He here places the kind of naturalistic research the present investigation has discussed in the context of a biting criticism of religious beliefs and practices. But is the thesis in Dawkins (2007) that CSR or other forms of naturalistic, Darwinian research on religion provide conclusive evidence against the truth of theism? Dawkins and Daniel Dennett (2007) are in recent debates often mentioned as possible advocates of the view that CSR debunks or even falsifies theism. That would make them adherents of the conflict-model described in chapter 1. For example, according to Aku Visala, both Dawkins and Dennett contend or at least tacitly imply that we can plausibly move from the naturalistic explanations CSR provides to atheism. The basis for this suggestion is that they both appeal to such research as part of their criticism of religion (Visala, 2011, 166). Visala is critical of the notion that naturalistic explanations of religion provide good grounds for atheism. We cannot in his view move plausibly from the notion that religion is explainable to the conclusion that religion is false or that religious beliefs are irrational (Visala, 2011, 160, cf. similar points by Leech and Visala, 2011, 301, and Clark and Barrett, 2011, 640-641).[115] This seems correct, but Dawkins does not argue that way.

I shall describe a basic and legitimate way to appeal to CSR in advocacy for atheism. Proceeding in this way is not necessarily indicative of a naturalistic orientation but is also

[115] Consider also the review of recent articles about debunking in chapter 1, section 5. In many of these, Dawkins (2007) and Daniel Dennett (2006) are mentioned as possible advocates for such a view.

available to those with an armchair orientation. Dawkins, who is a largely armchair oriented critic of religion, employs this strategy. From the perspective the current investigation has developed, this strategy is not the primary one to pursue. As I have shown, we can leave the armchair behind and go further in using cognitive science when criticising religion. Moreover, we need not attempt to demonstrate the truth of atheism to show that theistic beliefs in many cases lack justification. Let us nevertheless consider the role of CSR in the discussion Dawkins (2007) provides, to appreciate the legitimacy of this strategy.

To appreciate this strategy, let us look at some remarks Julian Baggini and Peter Fosl (2010) provide about how to persuasively defend a position. Baggini and Fosl here argue that those who offer a new thesis which departs strongly from widely held views that seem intuitively true should do two things: they should provide evidence for their thesis and bolster their case by offering an explanation of why alternative views are widely held despite being false. This is because when one departs radically from a commonly held belief, then both evidence to support that departure and some explanation for why people widely would be wrong about the matter is needed. Providing this explanation will make the case for the new thesis more persuasive to sceptics (Baggini and Fosl, 2010, 95). To illustrate, Baggini and Fosl (2010, 95-96) give the example of someone who defends the thesis that the earth is spherical in a context where the majority believe it to be flat. But let us instead, for our present purposes, envisage an altered case whereby someone defends heliocentrism in a context where the majority favours a geocentric view of the earth.[116] The geocentric model of the universe placed the earth at the centre and depicted the sun as orbiting earth. This view was widely held among laypeople, who depended on untutored intuitions and observations, and by clever thinkers like Aristotle and Ptolemy, who provided reasoned arguments in its defence. Moreover, the geocentric model Ptolemy carefully developed was instrumentally useful: it could (but not perfectly) predict the positions of the planets (see Lindberg, 2003). This was therefore no bizarre thesis irrationally held in the face of compelling counterevidence. Yet, it turned out to be wrong. This is also the case with theism according to the atheist advocate. Now to the illustration of why evidence for a thesis should be accompanied by an explanation of why the

[116] One might otherwise worry that theists are being compared to present day flat-earthers and that atheists are compared to present day spherical-earthers. But that is not what I am doing. Theists are not analogous to the former subjects. But neither are those who favoured the geocentric view in a historical context prior to the emergence of heliocentrism. In the historical 16th century context where initial versions of Heliocentrism were promoted, the geocentric alternative was widely held by laypeople and among the scientific elite. Moreover, the initial scepticism the elite showed towards the new model was not, as some might think, based on irrationality or superstition (see Lindberg, 2003). It also needs mentioning that the initial versions of heliocentrism offered by Copernicus and Galileo, of course, were worlds apart from the present-day view of the universe.

opposite thesis seems true despite being false. Baggini and Fosl (2010) point out that when someone offers a new thesis which departs radically from widely held beliefs, people will be sceptical. Evidence which strongly supports this thesis might not be immediately obvious or convincing. In our case, the intuitive pull of the notion that the earth is stationary and orbited by the sun can create problems for the advocates for the new thesis, supposing they aim to convince many rather than only a small elite. "The earth seems stationary. It is the sun that moves!", people might for example complain, and with good reason. It really seems to be so. Geocentric scientists can also devise arguments in support of these widely held beliefs. According to Baggini and Fosl (2010), it will therefore help the advocates for the new thesis if they can offer an explanation of why people would be widely wrong about this matter. This explanation should tell us why this wrong view seems so plausible. Let us call the explanation relevant to our case "geocentric psychology". Geocentric psychology needs to take account of the observations and intuitions that undergird the belief that the earth is stationary and orbited by the sun. It needs to tell us why earth will seem that way even though it orbits the sun. That explanation in itself does not, of course, establish the truth of the new thesis. We do not show that the earth orbits the sun by explaining why people widely, and with good reason even, believe that the sun orbits earth. However, by adding an explanation of why the latter nevertheless seems true, the primary case for the Heliocentric view becomes more plausible and persuasive. In short, backing up evidence for a new thesis with an account of why the alternative seems plausible or even obvious, makes for a persuasive defence (Baggini and Fosl, 2010, 95-96).

Let us now turn to Richard Dawkins's criticism of religion. The primary target in that criticism is what Dawkins calls the "God Hypothesis", the thesis that "there exists a superhuman, supernatural intelligence who deliberately designed and created the universe and everything in it, including us" (Dawkins, 2007, 52). The thesis Dawkins defends is that there almost certainly is no such intelligence or any other deity (Dawkins, 2007, 189). Is his defence of this atheistic thesis analogous to a case for a Heliocentric view which appeals only to geocentric psychology? Such an analogy should be present if the above-illustrated view about how he criticizes religion is true. As we saw above, Dawkins is depicted as someone who uses CSR to show that theism is false. But a brief look at the dialectic of his argument shows that this is not the case. The primary objection to the God hypothesis is that this hypothesis, which attempts to explain the universe by postulating a designer, faces an explanatory regress. The problem is that we do not explain the universe by postulating something that is more complex

and improbable than the universe itself, namely God[117] (Dawkins, 2007, 137-189). Dawkins also takes the (alleged) failure of central religious arguments to undermine the God hypothesis but this discussion plays a secondary role (Dawkins, 2007, 100-136). His criticism thus combines a direct attack on the God hypothesis with criticism of arguments in its favour. Importantly, this part of his argument is unrelated to CSR. Dawkins' criticism therefore seems more related to the evidential objection to theism (Ch. 4, section 2) than to debunking (in the sense we have used that term). What Dawkins is doing thus far in his argument therefore seems analogous to what the heliocentric advocate is doing when providing evidence for the view that the earth orbits the sun, before pursuing the additional issue of why the geocentric view is intuitively appealing. The analogy at play here, to be more specific, is between how the heliocentrism-advocate we envisage and Dawkins both bring evidence for their respective views, prior to using psychology to explain why people nevertheless would systematically hold the opposite views.[118]

Then, after concluding that there almost certainly is no supernatural intelligence of any kind, Dawkins (2007) appeals to CSR and the question of how one can apply Darwinian principles to religion. He here starts by asking why people would hold false religious beliefs and why religion is ubiquitous despite consisting of false claims. That is not the same as applying such principles to show that these beliefs are false. It seems more like what the Heliocentric advocate does when appealing to geocentric psychology to explain why the geocentric view would be widely held, despite being false. Moreover, it seems a legitimate question to ask given that one has already on independent grounds concluded that such beliefs indeed are false. We can relate what Dawkins here does to what scholars in CSR do. He for example here asks basically the same question as the one Stewart Guthrie asks (quoted in chapter 1, section 3): "why does religion exist if religion consists of a false assertion?" (Dawkins, 2007, 189, cf. Guthrie, 1995, 200). Dawkins also, in a playful and thought-provoking discussion, stresses that evolutionary speaking, the ubiquity of religion is a puzzle. The puzzle lies in the fact that religion tends to involve costly and often extravagant activities and that evolution is stingy. Such a large and complex phenomenon requires a large theory, and Darwinian principles are up for the challenge of providing the right investigative framework (Dawkins, 2007, 190-200, cf. Atran 2002, 4-7, 114-117 and Bulbulia and Sosis, 2011, 370).

[117] I.e. the superhuman intelligence specified above and therefore not the same as "God" the way I use that term.
[118] I am therefore not comparing Dawkins' arguments for atheism to Galileo's arguments for heliocentrism, for example. I am not here assessing the strength of Dawkins' argument but simply illustrating what role CSR plays in it.

Dawkins endorses and discusses the by-product thesis, here touching on the work of Pascal Boyer (Dawkins, 2007, 200, 218). Dawkins further suggests that we can combine the by-product thesis with his own theory of *memetics*, a theory which competes with the epidemiology-theory CSR-scholars (like Sperber and Boyer) apply (Dawkins, 2007, 222-234). Scholars wedded to epidemiology will probably object (see e.g. Sperber, 2000). We can have an interesting discussion about the relative merits of memetics and epidemiology. But the crucial point now is that the prospect of a Darwinian explanation of religion is not presented by Dawkins as a source of grounds to establish atheism. The explanatory regress problem and the alleged failure of arguments for the God hypothesis play that role. By contrast, the by-product thesis, memetics, and the Darwinian principles that Dawkins applies are discussed in separation from the attempt to establish atheism. Yet, these topics are clearly placed in a larger dialectical context whereby the primary goal is to attack religion. We can interpret this as an attempt to apply CSR to bolster the case for atheism, not an argument directly from CSR to atheism. Here is how this application can bolster a case for atheism.

It is initially puzzling that the earth orbits the sun given that it widely seems stationary while the sun really seems to move. It is also, one might argue, puzzling that atheism is true given that God and many other deities widely seem real. Given that human minds do not systematically make mistakes about most issues, why do they nevertheless make such mistakes by forming belief in deities? Usually and for good evolutionary reasons, things that are not real do not seem real. One might even suggest that atheism faces an explanatory burden. "God seems real!", the theist will object, and put an explanatory challenge to the atheist: "how can God seem real to so many people if he is not real"? This is a relevant question to put to an atheist. Encounters like this therefore make it relevant for the atheist to have some explanation of why people would widely hold the (allegedly) incorrect belief that God or other deities exist. Here the naturalistic explanations of religion CSR provides come into view as powerful auxiliary tools that aid the atheist in ways that resemble the aid geocentric psychology offered the heliocentric advocate. As I have argued, CSR can explain in a naturalist fashion how people become theists. Moreover, this explanation is a token of a more general explanation of how people develop belief in cognitively attractive deities in general. The advocate for atheism can, by integrating this naturalistic research on religion, provide a case that combines a defence of atheism with a psychology of (allegedly false) religious belief. Assuming that the defence of atheism, which is the primary part of this case, is convincing, the atheist advocate can then show both *that* theistic and other religious beliefs probably are wrong and explain *why* such

beliefs nevertheless would be widely entertained. The case of Dawkins (2007) shows that it is possible for a largely armchair oriented, evidentialist critic of religion to use CSR in this way.

The present investigation has not argued for atheism, and we have left the armchair for a naturalistic perspective in the philosophy of religion. This perspective sees empirical research in cognitive science as central in normative philosophical discussions, while also appreciating that there are arguments both for and against theism which cognitive science does not strongly bear on. This perspective enables us to go further than Dawkins in usage of CSR: we can use psychology to establish negative implications for theistic arguments and beliefs. This does not mean that we go further than Dawkins in his biting, often satirical criticism of religion. Instead, we go further in our usage of cognitive science when reaching normative philosophical conclusions concerning religion.

Bibliography

Adams, Robert. 1987. *The Virtue of Faith.* New York: Oxford University Press.

Alexander, Denis. 2014. *Creation or Evolution. Do we Have to Choose?* Monarch Books.

Alston, William P. 1996. *A Realist Conception of Truth.* Cornell University Press.

Alston, William P. 2005. *Beyond Justification. Dimensions of Epistemic Justification.* Cornell University Press.

Andersen, Marc. 2019. "Predictive Coding in Agency Detection". *Religion, Brain & Behavior* 9(1): 65-84. Doi:10.1080/2153599X.2017.1387170.

Aristotle. 2009. *The Nicomachean Ethics*. Oxford University Press.

Armstrong, Karen. 1999. *A History of God. From Abraham to the Present: the 4000-year Quest for God.* Vintage.

Atran, Scott. 2002. *In Gods We Trust. The Evolutionary Landscape of Religion.* Oxford University Press.

Atran, Scott and Ara Norenzayan. 2004. "Religion's Evolutionary Landscape: Counterintuition, commitment, compassion, communion". *Behavioral and Brain Sciences* 27: 713-770.

Austin, Alfredo Lopêz, and Leonardo Lopêz Lujan. 2008. "Aztec Human Sacrifice". In *The Aztec World*, edited by Elisabeth M. Brumfiel and Gary M. Feinman, 137-152. New York: Harry N. Abrams, Inc.

Baggini, Julian and Peter S. Fosl. 2010. *The Philosopher's Toolkit. A Compendium of Philosophical Concepts and Methods.* Wiley-Blackwell.

Banerjee, Konika, Omar S. Haque, and Elisabeth S. Spelke. 2013 "Melting Lizards and Crying Mailboxes: Children's Preferential Recall of Minimally Counterintuitive Concepts". *Cognitive Science* 37: 1251-1289.

Barrett, H. C. 2005. "Adaptations to Predators and Prey". In *The Handbook of Evolutionary Psychology*, edited by David Buss, 200-223. John Wiley & Sons, Inc.

Barrett, H. C. 2015. *The Shape of Thought. How Mental Adaptations Evolve*. Oxford University Press.

Barrett, Justin L. 1999. "Theological Correctness: Cognitive Constraint and the Study of Religion". *Method & Theory in the Study of Religion* 11: 325-339.

Barrett, Justin L. 2000. "Exploring the Natural Foundation of Religion". *Trends in Cognitive Sciences* 4(1): 29-34.

Barrett, Justin L. 2002. "Dumb Gods, Petitionary Prayer and the Cognitive Science of Religion". In *Current Approaches in the Cognitive Science of Religion*, edited by Illkka Pyysiäinen and Veikko Anttonen, 93-109. Continuum.

Barrett, Justin L. 2004. *Why Would Anyone Believe in God.* Altamira Press.

Barrett, Justin L. 2007. "Is the Spell Really Broken? Bio-psychological Explanations of Religion and Theistic Belief". *Theology and Science* 5(1): 57-72.

Barrett, Justin L. 2008a. “Coding and Quantifying Counterintuitiveness in Religious Concepts: Theoretical and Methodological Reflections”. *Method and Theory in the Study of Religion* 20: 308-338.

Barrett, Justin L. 2008b. “Why Santa Claus is not a God”. *Journal of Cognition and Culture.* 8:149-161.

Barrett, Justin L. and Frank Keil. 1996. “Conceptualizing a Nonnatural Entity: Anthropomorphism in God concepts”. *Cognitive Psychology* 31: 219-247.

Barrett, Justin L. and Melanie A. Nyhof. 2001. “Spreading Non-natural Concepts: The Role of Intuitive Conceptual Structures in Memory and Transmission of Cultural Materials”. *Journal of Cognition and Culture* 1(1), 69-100.

Barrett, Justin L. and Rebekah Richert. 2003. “Anthropomorphism or Preparedness? Exploring Children’s God Concepts”. *Review of Religious Research* 44(3): 300-312.

Barrett, Justin L. and Jonathan Lanman. 2008. “The Science of Religious Beliefs”. *Religion* 38(2): 109-124.

Bartlett, F. C. 1932. *Remembering.* Cambridge: Cambridge University Press.

Baumard, Nicolas and Pascal Boyer. 2013. “Explaining Moral Religions”. *Trends in Cognitive Sciences* 17(6): 272-280.

Baumard, Nicolas, Alexandre Hyafil, Ian Morris, and Pascal Boyer. 2015. “Increased Affluence Explains the Emergence of Ascetic Wisdoms and Moralizing Religions. *Current Biology* 25:10-15.

Bechtel, William and Cory D. Wright. 2009. “What is Psychological Explanation?” In *The Routledge Companion to Philosophy of Psychology,* edited by John Symons and Paco Calvo, 113-130. Routledge.

Berg, Jonathan. 1997. “How Could Ethics Depend on Religion?”. In *A Companion to Ethics,* edited by Peter Singer, 525-533. Blackwell Publishers.

Bergmann, Michael. 2010. “Epistemology of Religious Belief – Recent Developments”. In *A Companion to Epistemology,* editors Jonathan Dancy et al., 697-698. Wiley-Blackwell.

Blancke, Stefaan, Hans Henrik Hjermitslev, and Peter C. Kjærgaard. 2014. “Introduction: Creationism in Europe or European Creationism?” In *Creationism in Europe,* edited by Stefaan Blancke et al., 1-14. John Hopkins University Press.

Bloch, Maurice. 2002. “Are Religious Beliefs Counterintuitive?” in *Radical Interpretation in Religion,* edited by Nancy K. Frankenberry, 129-146. Cambridge University Press.

Bloom, Paul. 2004. *Descartes’ Baby: How the Science of Child Development Explains What Makes us Human.* Basic Books.

Bloom, Paul. 2010. “Religious Belief as an Evolutionary Accident” In *The Believing Primate. Scientific, Philosophical and Theological Reflections on the Origin of Religion,* edited by Jeffrey Schloss and Michael J. Murray, 118-127. Oxford University Press.

BonJour, Laurence. 2008. “Externalist Theories of Empirical Knowledge”. In *Epistemology. An*

Anthology, edited by Ernest Sosa et al, 363-378. Blackwell.

Bourget, David, and David J. Chalmers. 2014. "What do Philosophers Believe?" *Philosophical Studies* 170: 465-500.

Boyd, R. and Richerson, P. J. 1985. *Culture and the Evolutionary Process.* University of Chicago Press.

Boyd, R. 1988. "How to be a Moral Realist". In *Essays on Moral Realism,* edited by Sayre-McCord. Ithaca, NY: Cornell University Press.

Boyd, Robert and Joan B. Silk. 2015. *How Humans Evolved.* New York and London: W. W. Norton & Company.

Boyer, Pascal. 1994. *The Naturalness of Religious Ideas. A Cognitive Theory of Religion.* University of California Press.

Boyer, Pascal. 2000. "Functional Origins of Religious Concepts: Ontological and Strategic Selection in Evolved Minds". *The Journal of the Royal Anthropological Institute* 6(2): 195-214.

Boyer, Pascal. 2002. *Religion Explained. The Human Instincts that Fashion Gods, Spirits and Ancestors.* London: Vintage.

Boyer, Pascal. 2003. "Religious Thought and Behaviour as By-products of Brain Function". *Trends in Cognitive Sciences.* 7(3): 119-124.

Boyer, Pascal. 2018. *Minds Make Societies. How Cognition Explains the World Humans Create.* New Haven and London: Yale University Press.

Boyer, Pascal and Charles Ramble. 2001. "Cognitive Templates for Religious Concepts: Cross-cultural Evidence for Recall of Counter-intuitive Representations". *Cognitive Science* 25: 535-564.

Boyer, Pascal and H. C. Barrett. 2005. "Domain Specificity and Intuitive Ontology". In *The Handbook of Evolutionary Psychology,* edited by David Buss, 96-118. John Wiley & Sons, Inc.

Boyer, Pascal, and Pierre Lienard. 2006. "Why Ritualized Behavior? Precaution Systems and Action Parsing in Developmental, Pathological, and Cultural Rituals". *Behavioral and Brain Sciences* 29: 595-650.

Buchanan, Allen, and Russell Powell. 2018. *The Evolution of Moral Progress. A Biocultural Theory.* Oxford University Press.

Braddock, Matthew. 2016. "Debunking Arguments and the Cognitive Science of Religion". *Theology and Science* 14(3): 268-287.

Brink, D. 1989. *Moral Realism and the Foundations of Ethics.* Cambridge: Cambridge University Press.

Bulbulia, Joseph and Richard Sosis. 2011. "Signalling Theory and the Evolution of Religious Cooperation". *Religion* 41(3): 363-388.

Buller, David. 2006. *Adapting Minds. Evolutionary Psychology and the Persistent Quest for Human Nature.* MIT Press.

Buss, David. 2015. *Evolutionary Psychology. The New Science of the Mind.* 5th edition. London and New York: Routledge.

Caldwell-Harris, Catherine, C. Murphy, C. F. Velazquez, P. McNamara. 2011. "Religious Belief Systems of Persons with High Functioning Autism". *Proceedings of the Annual Meeting of the Cognitive Science Society.* 33. Online Publication: https://escholarship.org/uc/item/6zh3j3pr

Callahan, Laura Frances, and Timothy O'Connor. 2014. "Introduction". In *Religious Faith and Intellectual Virtue,* edited by L. F. Callahan and T. O'Connor, 1-26. Oxford University Press.

Clark, Kelly James and Justin L. Barrett. 2010. "Reformed Epistemology and the Cognitive Science of Religion". *Faith and Philosophy* 27(2): 174-189.

Clark, Kelly James and Justin L. Barrett. 2011. "Reidian Religious Epistemology and the Cognitive Science of Religion". *Journal of the American Academy of Religion* 79(3): 639-675.

Clark, Kelly James and Dani Rabinowitz. 2018. "Knowledge and the Objection to Religious Belief from Cognitive Science". In *The Roots of Religion. Exploring the Cognitive Science of Religion,* edited by Roger Trigger and Justin L. Barrett, 113-125. London and New York: Routledge.

Conee, Earl and Richard Feldman. 2011. *Evidentialism. Essays in Epistemology.* Oxford: Clarendon Press.

Copan, Paul. 2008. "The Moral Argument". In *Philosophy of Religion: Classic and Contemporary Issues,* edited by P. Copan and C. Meister, 127-141. Malden, MA: Blackwell.

Cosmides, Leda, John Tooby and Jerome H. Barkow. 1992. "Introduction. Evolutionary Psychology and Conceptual Integration". In *The Adapted Mind. Evolutionary Psychology and the Generation of Culture,* edited by J. H. Barkow, L. Cosmides & J. Tooby, 3-15. New York: Oxford University Press.

Cosmides, Leda and John Tooby. 2005. "Neurocognitive Adaptations Designed for Social Exchange". In *The Handbook of Evolutionary Psychology,* edited by David Buss, 584-627. John Wiley & Sons, Inc.

Coyne, Jerry. 2015. *Faith Versus Fact. Why Science and Religion Are Incompatible.* Viking.

Craig, William L. and J. P. Moreland. 2009. "Introduction". In *Blackwell Companion to Natural Theology,* edited by W. L. Craig and J. P. Moreland, ix-xiii. Wiley-Blackwell.

Czachesz, Istvan. 2017. *Cognitive Science and the New Testament. A New Approach to Early Christian Research.* Oxford University Press.

Darwall, Stephen. 1998. *Philosophical Ethics.* Westview Press.

Dawkins, Richard 1998. *Unweaving the Rainbow. Science, Delusion and The Appetite for Wonder*. The Penguin Press.

Dawkins, Richard. 2007. *The God Delusion.* Transworld Publishers.

Dawkins, Richard. 2016. *The Extended Phenotype. The Long Reach of the Gene*. Oxford University Press.

De Cruz, Helen and Johan De Smedt. 2015. *A Natural History of Natural Theology. The Cognitive Science of Theology and Philosophy of Religion.* MIT Press.

De Cruz, Helen and Johan De Smedt. 2016. “How Do Philosophers Evaluate Natural Theological Arguments?” In *Advances in Religion, Cognitive Science, and Experimental Philosophy,* edited by Helen De Cruz and Ryan Nichols, 119-142. Bloomsbury.

De Cruz, Helen and Ryan Nichols. 2016. “Introduction”. In *Advances in Religion, Cognitive Science, and Experimental Philosophy,* edited by Helen De Cruz and Ryan Nichols, 1-17. Bloomsbury.

Dennett, Daniel. 2006. *Breaking the Spell. Religion as a Natural Phenomenon.* Penguin Books.

Dougherty, Trent. 2011. “Introduction”. In *Evidentialism and its Discontents,* edited by Trent Dougherty, 1-14. Oxford University Press.

Duntley, Joshua D. 2005. “Adaptations to Dangers from Humans“. In *The Handbook of Evolutionary Psychology,* edited by David Buss, 224-249. John Wiley & Sons, Inc.

Elster, Jon. 2007. *Explaining Social Behaviour. More Nuts and Bolts for the Social Sciences.* Cambridge University Press.

Evans, Stephen C. 2018. “Moral Arguments for the Existence of God”. *The Stanford Encyclopedia of Philosophy* (Fall 2018 edition), edited by Edward N. Zalta. Online Publication. URL = <https://plato.stanford.edu/archives/fall2018/entries/moral-arguments-god/>.

Fitzpatrick, William J. 2014. “Why There is No Darwinian Dilemma for Ethical Realism.” In *Challenges to Moral and Religious Belief. Disagreement and Evolution,* edited by Michael Bergmann and Patrick Kain, 237-255. Oxford University Press.

Flood, Gavin. 2004. *The Ascetic Self. Subjectivity, Memory and Tradition.* Cambridge University Press.

Fodor, Jerry A. *The Modularity of Mind.* MIT Press, 1987.

Forstmann, Mattias, and Pascal Burgmer. 2015. “Adults are Intuitive Mind-Body Dualists”. *Journal of Experimental Psychology* 144(1): 222-235.

Gelman, Susan A., John D. Doley, Gail M. Gottfried. 1994. “Essentialist Beliefs in Children: The Acquisition of Concepts and Theories”. In *Mapping The Mind. Domain Specificity in Cognition and Culture,* edited by L. A. Hirschfeld and S. A. Gelman, 341-365. Cambridge University Press.

Gervais, Will M. and Joseph Henrich. 2010. “The Zeus Problem: Why Representational Content Biases Cannot Explain Faith in Gods”. *Journal of Cognition and Culture* 10: 383-389.

Gervais, Will M., Aiyana K. Willard, Ara Norenzayan and Joseph Henrich. 2011. “The Cultural Transmission of Faith. Why Innate Intuitions are Necessary, but Insufficient, to Explain Religious Belief”. *Religion* 41(3): 389-410.

Gervais, Will M. and Ara Norenzayan. 2012. “Like a Camera in the Sky? Thinking about God Increases Public Self-Awareness and Socially Desirable Responding”. *Journal of Experimental Social Psychology* 48: 298-302.

Godfrey-Smith, Peter. 2003. *Theory and Reality. An Introduction to the Philosophy of Science.* University of Chicago Press.

Goldman, Alvin. 1976. “Discrimination and Perceptual Knowledge”. *The Journal of Philosophy* 73 (20): 771-791.

Goldman, Alvin. 1986. *Epistemology and Cognition.* Cambridge and London: Harvard University Press.

Goldman, Alvin. 1988. “Strong and Weak Justification”. *Philosophical Perspectives.* 2: 51-69.

Goldman, Alvin. 2015. *Reliabilism and Contemporary Epistemology. Essays.* Oxford University Press.

Goldman, Alvin and Matthew McGrath. 2015. *Epistemology. A Contemporary Introduction.* Oxford University Press.

Goodnick, Liz. 2016. “A *De jure* Criticism of Theism”. *Open Theology* 2: 23-33.

Gottschling, Vera. 2009. “Modularity”. In *The Routledge Companion to Philosophy of Psychology*, edited by John Symons and Paco Calvo, 296-321. Routledge.

Gould, Stephen J. 2002. *Rocks of Ages. Science and Religion in the Fullness of Time.* Vintage.

Grant, Robert M., David Tracy. 2005. *A Short History of the Interpretation of the Bible.* Fortress Press.

Greco, John. 1993. “Virtues and Vices of Virtue Epistemology”. *Canadian Journal of Philosophy* 23(3): 413-432.

Greco, John. 2007. *Putting Skeptics in their Place. The Nature of Skeptical Arguments and Their Role in Philosophical Inquiry.* Cambridge University Press.

Greco, John. 2010. *Achieving Knowledge. A Virtue-Theoretic Account of Epistemic Normativity.* Cambridge University Press.

Groome, David, H. Dewart, A. Esgate, K. Gurney, R. Kemp, N. Towell. 1999. *An Introduction to Cognitive Psychology. Processes and Disorders.* Psychology Press Ltd, Publishers.

Guthrie, Stewart. 1980. “A Cognitive Theory of Religion”. *Current Anthropology* 21(2): 181-203.

Guthrie, Stewart. 1995. *Faces in the Clouds. A New Theory of Religion.* New York and Oxford: Oxford University Press.

Guthrie, Stewart. 2002. “Animal Animism: Evolutionary Roots of Religious Cognition”. In *Current Approaches in the Cognitive Science of Religion,* edited by Illka Pyysiäinen and Veikko Anttonen, 38-67. London: Continuum.

Guthrie, Stewart. 2017. “Religion Explained? Some Variants of Cognitive Theory”. In *Religion Explained? The Cognitive Science of Religion After Twenty-five Years,* edited by Luther H. Martin and Donald Wiebe. Bloomsbury.

Hare, John. 2004. “Is there an Evolutionary Foundation for Human Morality? In *Evolution and Ethics: Human Morality in Biological and Religious Perspective,* edited by P. Clayton and J. Schloss, 187-203. Wm. B. Eerdmans Publishing Company.

Richert, Rebekah A., and Paul L. Harris. 2008. “Dualism Revisited: Body *vs.* Mind *vs.* Soul”. *Journal of Cognition and Culture* 8: 99-115.

Haselton, Martie G., Daniel Nettle and Paul W. Andrews. 2005. “The Evolution of Cognitive Bias”. In *The Handbook of Evolutionary Psychology,* edited by David Buss, 724-746. John Wiley & Sons, Inc.

Haught, John F. 2011. “Darwin, Design, and Divine Providence”. In *Debating Design. From Darwin to DNA*, edited by William A. Dembski and Michael Ruse, 229-245. Cambridge University Press.

Hirschfeld, Lawrence A. and Susan A. Gelman. 1994. “Toward a Topography of Mind: An Introduction to Domain Specificity”. In *Mapping The Mind. Domain Specificity in Cognition and Culture*, edited by L. A. Hirschfeld and S. A. Gelman, 3-35. Cambridge University Press.

Hood, Ralph, W., Peter C. Hill and Bernard Spilka 2009. *Introduction to Psychology of Religion. An Empirical Approach*. London and New York: Guilford Press.

Johnson, Dominic. 2016. *God is Watching You. How The Fear of God Makes us Human.* Oxford University Press.

Johnson, Dominic, and Jesse Bering. 2010. “Hand of God, Mind of Man: Punishment and Cognition in the Evolution of Cooperation”. In *The Believing Primate. Scientific, Philosophical and Theological Reflections on the Origin of Religion,* edited by Jeffrey Schloss and Michael J. Murray, 26-43. Oxford University Press.

Jones, James W. 2016. *Can Science Explain Religion? The Cognitive Science Debate.* Oxford University Press.

Jong, Jonathan and Aku Visala. 2014. “Evolutionary debunking-arguments Reconsidered”. *International Journal for Philosophy of Religion* 76: 243-258.

Joyce, Richard. 2006. *The Evolution of Morality.* Cambridge MA and London: MIT Press.

Kahane, Guy. 2011. “Evolutionary Debunking Arguments”. *Noûs* 45(1): 103-125.

Kelemen, Deborah. 2004. “Are Children “Intuitive Theists”? Reasoning About Purpose and Design in Nature”. *Psychological Science* 15: 295-301.

Kirkpatrick, Lee. 2005. *Attachment, Evolution, and the Psychology of Religion.* New York and London: Guilford Press.

Krebs, Dennis. 2005. “The Evolution of Morality”. In *The Handbook of Evolutionary Psychology*, edited by David Buss, 747-771. John Wiley & Sons, Inc.

Kurzban, Robert and Steven Neuberg. 2005. “Managing Ingroup and Outgroup Relationships”. In *The Handbook of Evolutionary Psychology*, edited by David Buss, 653-675. John Wiley & Sons, Inc.

Laland, Kevin N. 2017. *Darwin’s Unfinished Symphony. How Culture Made the Human Mind*. Princeton University Press.

Larson, Edward J. 2003. “The Scopes Trial in History and Legend”. In *When Science & Christianity Meet,* edited by David C. Lindberg and Ronald L. Numbers, 245-264. University of Chicago Press.

Lawson, E. T., and R. McCauley. 1990. *Rethinking Religion: Connecting Cognition and Culture.* Cambridge: Cambridge University Press.

Levy, Gabriel. 2014. *Judaic Technologies of the Word. A Cognitive Analysis of Jewish Cultural Formation.* Routledge.

Lewis, C. S. 1952. *Mere Christianity*. The Macmillan Company.

Leech, David and Aku Visala 2011. "The Cognitive Science of Religion: A Modified Theist Response". *Religious Studies* 47: 301-316.

Lindberg, David C. 2003. "Galileo, the Church, and the Cosmos". In *When Science and Christianity Meet,* edited by David C. Lindberg and Ronald L. Numbers, 33-60. Chicago and London: University of Chicago Press.

Linville, Mark D. 2009. "The Moral Argument". In *Blackwell Companion to Natural Theology,* edited by W. L. Craig and J. P. Moreland, 391-448. Wiley-Blackwell.

Luhrmann, T. M. 2012. *When God Talks Back. Understanding the American Evangelical Relationship with God.* New York: Vintage Books.

Lyons, Jack C. 2009. *Perception and Basic Beliefs. Zombies, Modules, and the Problem of the External World.* Oxford University Press.

Mackie, John. 1974. *The Cement of the Universe. A Study In Causation.* Oxford University Press.

Maij, David L. R, Hein T. Van Schie, and Michiel Van Elk. 2019. "The Boundary Conditions of the Hypersensitive Agency Detection Device: An Empirical Investigation of Agency Detection in Threatening Situations". *Religion, Brain & Behaviour* 9 (1): 23-51.

Harmon-Vukic, Mary, M. Afzal Upal, and Kelly J. Sheehan. 2012. "Understanding the Memory Advantage of Counterintuitive Concepts". *Religion, Brain & Behavior* 2(2): 121-139.

McCauley, Robert N. 2013. *Why Religion is Natural and Science is Not.* Oxford University Press.

McCauley, Robert N. 2017a. *Philosophical Foundations of the Cognitive Science of Religion. A Head Start.* Bloomsbury.

McCauley, Robert N. 2017b. "Twenty-Five Years In: Landmark Empirical Findings in the Cognitive Science of Religion". In *Religion Explained. The Cognitive Science of Religion After Twenty-Five Years,* edited by Luther H. Martin and Donald Wiebe, 17-55. Bloomsbury.

McCutcheon, Russell T. 2010. "Will Your Cognitive Anchor Hold in the Storms of Culture?" *Journal of the American Academy of Religion* 78 (4): 1182-1193.

Mercier, Hugo and Dan Sperber. 2017. *The Enigma of Reason. A New Theory of Human Understanding.* Allen Lane.

Miller, Alexander. 2013. *An Introduction to Contemporary Metaethics.* Polity.

Moore, Andrew, and Michael Scott. 2007. "Introduction". In *Realism and Religion. Philosophical and Theological Perspectives,* edited by Andrew Moore and Michael Scott, 1-9. Ashgate.

Murray, Michael J. 2010. "Scientific Explanations of Religion and the Justification of Religious Belief". In *The Believing Primate. Scientific, Philosophical and Theological Reflections on the Origin of Religion,* edited by Jeffrey Schloss and Michael J. Murray, 168-178. Oxford University Press.

Myers, David G. 2002. *Social Psychology.* 7th ed. McGraw Hill.

Mysterud, Iver. 2005. *Mennesket og moderne evolusjonsteori*. Gyldendal Akademisk.

Nagasawa, Yujin. 2018. “The Problem of Evil for Atheists”. In *The Problem of Evil: Eight Views in Dialogue*, edited by N. N. Trakakis, 169-175. Oxford University Press.

Nicholson, Hugh. 2016. *The Spirit of Contradiction in Christianity and Buddhism.* Oxford University Press.

Norenzayan, Ara. 2013. *Big Gods. How Religion Transformed Cooperation and Conflict*. Princeton University Press.

Norezayan, Ara, Will M. Gervais. 2013. “The Origins of Religious Disbelief”. *Trends in Cognitive Sciences* 17(1): 20-25.

Norenzayan, A., A. F. Shariff, W. M. Gervais, A. K. Willard, R. A. McNamara, E. Slingerland, and J. Henrich. 2016. “The Cultural Evolution of Prosocial Religion”. *Behavioral and Brain Sciences* 39: 1-27.

Numbers, Ronald I. 2011. “Scientific Creationism and Intelligent Design”. In *Cambridge Companion to Science and Religion*, edited by Peter Harrison, 127-147. Cambridge University Press.

Numbers, Ronald I. 2014. “Foreword”. In *Creationism in Europe*, edited by Stefaan Blancke, Hans Henrik Hjermitslev, and Peter C. Kjærgaard, vii-xvi. John Hopkins University Press.

Paul II, John. 1997. “Message to the Pontifical Academy of Sciences”. *The Quarterly Review of Biology* 76(4): 381-383.

Peterson, Michael, William Hasker, Bruce Reichenbach and David Basinger. 1997. *Förnuft och religiös tro. En inledning til religionsfilosofin.* Bokförlaget Nye Doxa.

Pew Research Center. 2014. “Worldwide, Many See Belief in God as Essential to Morality”. Pew Research Center. Accessed April 21, 2019. https://www.pewglobal.org/2014/03/13/worldwide-many-see-belief-in-god-as-essential-to-morality/.

Plantinga, Alvin. 1983. “Reason and Belief in God”. In *Faith and Rationality. Reason and Belief in God*, edited by Alvin Plantinga and Nicholas Wolterstorff, 16-93. Notre Dame and London: University of Notre Dame Press.

Plantinga, Alvin. 1993. *Warrant and Proper Function.* New York: Oxford University Press.

Plantinga, Alvin. 2000. *Warranted Christian Belief.* New York: Oxford University Press.

Porubanova-Norquist, Michaela, Daniel Joel Shaw, and Dimitris Xygalatas. 2013. “Minimal-Counterintuitiveness Revisited: Effects of Cultural and Ontological Violations on Concept Memorability”. *Journal for the Cognitive Science of Religion* 1(2): 181-192.

Prichard, Duncan. 2010. “Sensitivity and Safety”. In *A Companion to Epistemology*, edited by Jonathan Dancy et al., 732-736. Wiley-Blackwell.

Prichard, Duncan. 2011. “Evidentialism, Internalism, Disjunctivism”. In *Evidentialism and its Discontents*, edited by Trent Dougherty, 235-253. Oxford University Press.

Proudfoot, Wayne. 1985. *Religious Experience.* Berkeley: University of California Press.

Purzycki, Benjamin G., Daniel L. Finkel, John Shaver, Nathan Wales Adam B. Cohen, and Richard Sosis. 2012. "What Does God Know? Supernatural Agents' Access to Socially Strategic and Non-Strategic Information." *Cognitive Science* 36: 846-869.

Purzycki, Benjamin G. 2013. "The Minds of Gods: A Comparative Study of Supernatural Agency". *Cognition* 129: 163-179.

Purzycki, Benjamin G. and Aiyana Willard. 2015. "MCI theory: A Critical Discussion". *Religion, Brain & Behavior.* Online Publication. DOI: 10.1080/2153599X.2015.1024915.

Pyysiäinen, Ilkka. 2009. *Supernatural Agents. Why We Believe in Souls, Gods, and Buddhas.* Oxford University Press.

Pyysiäinen, Ilkka, and Mark Hauser. 2009. "The Origins of Religion: Evolved Adaptation or By-product?" *Trends in Cognitive Sciences* 14(3): 104-109

Railton, P. 2003. *Facts, Values and Norms.* Cambridge: Cambridge University Press.

Randolph-Seng, B. and Michael E. Nielsen. 2007. "Honesty: One Effect of Primed Religious Representations". *The International Journal for the Psychology of Religion.* 17(4): 303-315.

Richardson, Robert C. 2007. *Evolutionary Psychology as Maladapted Psychology.* MIT Press.

Ritchie, Jack. 2008. *Understanding Naturalism.* Acumen Publishing.

Roberts, Robert C, and W. Jay Wood. 2014. *Intellectual Virtues. An Essay in Regulative Epistemology.* Oxford: Clarendon Press.

Rossano, Matt. 2009. "The African Interregnum: The "Where", "When", and "Why" of the Evolution of Religion". In *The Biological Evolution of Religious Mind and Behaviour*, edited by Eckart Voland and Wulf Schiefenhövel, 127-141. Springer-Verlag.

Ruse, Michael, E. O. Wilson. 1986. "Moral Philosophy as Applied Science". *Philosophy* 61(236): 173-192.

Ruse, Michael. 2011. "Atheism, Naturalism and Science: Three in One?" In *Cambridge Companion to Science and Religion,* edited by Peter Harrison, 229-243. Cambridge University Press.

Scerri, Eleanor M. L., M. G. Thomas, A. Manica, P. Gunz, J. T. Stock, C. Stringer, M. Grove, H. S. Groucutt, A. Timmermann, G. P. Rightmire, F. d'Errico, C. A. Tryon, N. A. Drake, A. S. Brooks, R. W. Dennell, R. Durbin, B. M. Henn, J. Lee-Thorp, P. deMenocal, M. D. Petraglia, J. C. Thompson, A. Scally, and L. Chikhi. "Did Our Species Evolve in Subdivided Populations Across Africa, and Why Does it Matter?" *Trends in Ecology and Evolution* 33(8): 582-594.

Schjoedt, Uffe. 2009. "The Religious Brain: A General Introduction to the Experimental Neuroscience of Religion". *Method and Theory in the Study of Religion.* 21: 310-339.

Schmidt, Klaus. 2010. "Göbekli Tepe – The Stone Age Sanctuaries. New Results of Ongoing Excavations with a Special Focus on Sculptures and high Reliefs". *Documenta Praehistorica* 37: 239-256. Online Publication. https://doi.org/10.4312/dp.37.21

Schloss, Jeffrey. 2010. “Introduction: Evolutionary Theories of Religion”. In *The Believing Primate. Scientific, Philosophical, and Theological Reflections on the Origin of Religion,* edited by Jeffrey Schloss & Michael J. Murray, 1-25. Oxford University Press.

Shafer-Landau. 2003. *Moral Realism.* Oxford: Oxford University Press.

Shafer-Landau. 2012. “Evolutionary Debunking, Moral Realism and Moral Knowledge.” *Journal of Ethics and Social Philosophy* 7(1): 1-37.

Slingerland, Edvard and Joseph Bulbulia. 2011. “Introductory essay: Evolutionary science and the study of religion”. *Religion* 41(3): 307-328.

Slone, Jason D., and William W. McCorcle Jr. 2019. *The Cognitive Science of Religion. A Methodological Introduction to Key Empirical Studies.* Bloomsbury Academic.

Sosa, Ernest. 2007. *A Virtue Epistemology. Apt Belief and Reflective Knowledge, Volume 1.* Oxford: Clarendon Press.

Sosa, Ernest. 2017. *Epistemology.* Princeton and Oxford: Princeton University Press.

Sosis, Richard. 2009. “The Adaptationist-Byproduct Debate on the Evolution of Religion: Five Misunderstandings of the Adaptationist Program. *Journal of Cognition and Culture.* 9: 315-332.

Spelke, Elisabeth, and Katherine D. Kinzler. 2007. “Core Knowledge”. *Developmental Science* 10(1): 89-96

Sperber, Dan. 1996. *Explaining Culture. A Naturalistic Approach.* Blackwell.

Sperber, Dan. 2000. “An objection to the Memetic Approach to Culture”. In *Darwinizing Culture. The Status of Memetics as a Science,* edited by Robert Aunger, 163-173. Oxford University Press.

Stenmark, Mikael. 2010. “Ways of Relating Science and Religion”. In *Cambridge Companion to Science and Religion,* edited by Peter Harrison, 278-295. Cambridge University Press.

Sudduth, Michael. 2009. *The Reformed Objection to Natural Theology.* Ashgate.

Swinburne, Richard. 2004. *The Existence of God.* Oxford: Clarendon Press.

Taliaferro, Charles. 2004. *Contemporary Philosophy of Religion.* Blackwell Publishers.

Taliaferro, Charles, 2010. “Explaining Religious Experience”. In *The Believing Primate. Scientific, Philosophical, and Theological Reflections on the Origin of Religion,* edited by Jeffrey Schloss & Michael J. Murray, 200-214. Oxford University Press.

Teehan, John. 2010. *In the Name of God. The Evolutionary Origins of Religious Ethics and Violence.* Wiley-Blackwell.

Teehan, John. 2016a. “Cognitive Science, Evil, and God”. In *Advances in Religion, Cognitive Science, and Experimental Philosophy*, edited by Helen De Cruz and Ryan Nichols, 39-60. Bloomsbury.

Teehan, John. 2016b. “Religion and Morality: The Evolution of the Cognitive Nexus”. In *The Oxford Handbook of Evolutionary Psychology and Religion,* edited by James R. Liddle and Todd K. Shackelford. Online Publication. DOI: 10.1093/oxfordhb/9780199397747.013.11.

Teehan, John. 2018. “Cognitive Science and the Limits of Theology”. In *The Roots of Religion. Exploring the Cognitive Science of Religion,* edited by Roger Trigg and Justin L. Barrett, 167-187. Routledge.

Thurow, Joshua. 2013. “Does cognitive science show belief in god to be irrational? The epistemic consequences of the cognitive science of religion”. *International Journal for Philosophy of Religion* 74: 77-98.

Tinbergen, Nico. 1963. “On Aims and Methods of Ethology”. *Zeitschrift für Tierpsychologie* 20: 410-433.

Tomasello, Michael. 2016. *A Natural History of Human Morality.* Harvard University Press.

Turiel, Elliot. 2008. “The Development of Morality”. In *Child and Adolescent Development. An Advanced Course,* edited by William Damon and Richard M. Lerner, 473-514. John Wiley & Sons, Inc.

Uro, Risto. 2016. *Ritual and Christian Beginnings. A Socio-Cognitive Analysis.* Oxford University Press.

Van Leeuwen, Neil and Michiel Van Elk. 2018. “Seeking the Supernatural: The Interactive Religious Experience Model”. *Religion, Brain & Behaviour*. Online Publication: DOI: 10.1080/2153599X.2018.1453529.

Van Eyghen, Hans, 2018. *Arguing From Cognitive Science of Religion: Cognitive Science of Religion and its Epistemic Implications.* PhD dissertation. Vrije Universiteit Amsterdam.

Visala, Aku. 2011. *Naturalism, Naturalism, Theism and the Cognitive Study of Religion.* Ashgate.

Watts, J., S. J. Greenhill, Q. D. Atkinson, T. E. Curie, J. Bulbulia, R. D. Gray. 2015. “Broad Supernatural Punishment but not Moralizing High Gods Precede the Evolution of Political Complexity in Austronesia.” *Proceedings of the Royal Society* B 282 20142556. Online Publication. http://dx.doi.org/10.1098/rspb.2014.2556.

Whitehouse, Harvey. 2000. *Arguments and Icons. Divergent Modes of Religiosity.* Oxford University Press.

Whitehouse, Harvey. 2004. *Modes of Religiosity: A Cognitive Theory of Religious Transmission: A Cognitive Theory of Religious Transmission.* AltaMira Press.

Whitehouse, Harvey, P. Franscois, P. E. Savage, T. E. Currie, K. C. Feeney, E. Cioni, R. Purcell, R. M. Ross, J. Larson, J. Baines, B. Haar, A. Covey, and P. Turchin. 2019. “Complex Societies Precede Moralizing Gods Throughout World History”. *Nature.* 586: 226-229. Online Publication. https://doi.org/10.1038/s41586-019-1043-4.

Whiten, Andrew. 2013. “Culture and the Evolution of Interconnected Minds.” In *Understanding Other Minds. Perspectives from Developmental Social Neuroscience*, edited by S. Baron-Cohen, Helen Tager-Flusberg, and Michael V. Lombardo, 431-447. Oxford University Press.

Wilkins, John S. and Paul E. Griffiths. 2012. “Evolutionary Debunking Arguments in Three Domains: Fact, Value, and Religion”. In *A New Science of Religion*, edited by James MacLaurin and Greg Dawes, 133-146. Routledge.

Willard, Aiyana K., and Ara Norenzayan. 2013. “Cognitive Biases Explain Religious Belief, Paranormal Belief, and Belief in Life’s Purpose. *Cognition.* 129: 379-391.

Wilson, Deidre and Dan Sperber. 2004. “Relevance Theory”. In *The Handbook of Pragmatics*, edited by Laurence R. Horn and Gregory Ward, 607-632. Blackwell Publishing.

Wilson, David Sloan. 2002a. *Darwin’s Cathedral. Evolution, Religion, and the Nature of Society.* University of Chicago Press.

Wilson, Timothy D. 2002b. *Strangers to Ourselves. Discovering the Adaptive Unconscious.* Cambridge and London: Harvard University Press.

Worrall, John 2005. “Science Discredits Religion”. In *Contemporary Debates in Philosophy of Religion*, edited by Michael L. Peterson and Raymond J. Vanarragon, 59-90. Blackwell Publishing.

Wright, Robert. 2004. *The Evolution of God. The Origins of Our Beliefs.* Little Brown and Company.

Zagzebski, Linda and Abrol Fairweather. 2001. “Introduction”. in *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, edited by A. Fairweather and L. Zagzebski, 3-14. Oxford University Press.

Zagzebski, Linda. 2008. “*Virtues of the Mind.* Selections”. In *Epistemology. An Anthology*, edited by Ernest Sosa et al., 442-453. Blackwell.

Zhong, Chen-Bo, Vanessa K. Bohns, and Francesca Gino. 2010. “Good Lamps are the Best Police: Darkness Increases Dishonesty and Self-Interested Behavior”. *Psychological Science* 21(3): 311-314.