

Values in science

The role of cognitive
and non-cognitive values in
Evolutionary Psychology

Silvia Ivani

Values in Science

The role of cognitive and non-cognitive values
in Evolutionary Psychology

Values in science

**The role of cognitive and non-cognitive values in
Evolutionary Psychology**

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University,
op gezag van de rector magnificus, prof. dr. K. Sijtsma, in het openbaar te
verdedigen ten overstaan van een door het college voor promoties aangewezen
commissie aan Tilburg University
op dinsdag 26 mei 2020 om 10.00 uur

door
Silvia Ivani,
geboren te La Spezia, Italië

promotor:	prof. dr. J.M. Sprenger
copromotor:	dr. M. Colombo
leden promotiecommissie:	prof. dr. C. Dutilh Novaes prof. dr. M.M.S.K. Sie prof. dr. H.W. De Regt dr. L. Henderson

Printed by: Proefschriftmaken | Proefschriftmaken.nl

© 2020 Silvia Ivani, The Netherlands. All rights reserved. No parts of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission of the author. Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd, in enige vorm of op enige wijze, zonder voorafgaande schriftelijke toestemming van de auteur.

To Sara

Abstract

Should scientists value simple theories? Is fruitfulness an important criterion to assess scientific theories? What is the role of moral, social, and political values in science? In recent years, there has been an increasing interest among philosophers of science in studying how cognitive and non-cognitive values influence and should influence the assessment and comparison of scientific theories. While cognitive values (e.g. simplicity and fruitfulness) are features of scientific theories that are indicative of the truth or empirical adequacy of theories, non-cognitive values are moral, political, social, and economic values. Understanding the roles of values in science is a particularly urgent issue. Clarifying the importance of cognitive values is important in order to be able to make accurate comparisons of scientific theories. Understanding the influence of non-cognitive values on science is crucial because moral, social, and political values are involved in many stages of research, such as decisions on methodologies and allocation of funds. Since these decisions affect all the members of the society (scientists, non-experts, and political institutions), understanding the impact of non-cognitive values on these choices is of primary importance. Although a large body of literature has investigated the roles of cognitive and non-cognitive values in theory appraisal, several questions remain to be analysed. In this dissertation, I address four questions by using two main methods, namely the analysis of case studies from Evolutionary Psychology and the experimental method. First, philosophers have clarified the importance and roles in theory appraisal of some values such as simplicity, whereas little or no attention has been paid to other values, such as fruitfulness. In the second chapter of my dissertation, I start fill this gap by formulating a clear explication and a strategy to be employed in theory appraisal for the value of fruitfulness. Moreover, I use my account to assess the fruitfulness of Evolutionary Psychology. Second, some philosophers of science argue that cognitive values are desirable and relevant to the assessment of theories because they are indicative of the truth or empirical adequacy of theories. However, why is it so? In the third chapter of my dissertation, I develop a context-sensitive approach to values. I argue that in order to understand the ground for the desirability of cognitive values and make an accurate appraisal of theories, we have to consider specific factors of the context in which theories are assessed, such as the availability of methodologies and the way cognitive values are interrelated to each other in that context. Third, philosophers have traditionally argued that the influence of non-cognitive values on scientific reasoning threatens the epistemic authority of science. In the third chapter, I challenge this view and argue that some non-cognitive values can play a cognitive role in science, i.e., they can be epistemically beneficial to the assessment of scientific theories. On the basis of a case study (the account of human mating in Evolutionary

Psychology), I argue that feminist values can positively contribute to theory appraisal in various ways, such as by raising sensitivity to evidence that was neglected because of gender bias. Fourth, some philosophers claim that non-cognitive values play a legitimate role in cases of inductive risk, namely cases in which scientists may wrongly assess scientific hypotheses (e.g. accepting a hypothesis that should be rejected) because of uncertainty. Mistakes can have consequences that can be morally or economically undesirable and non-cognitive values, some philosophers argue, provide the standards to evaluate and compare these possible consequences. However, little is known on how - specifically - non-cognitive values influence this evaluation. To address this issue, I conduct an experimental study clarifying how personal features, political values, and specific aspects of a risk (e.g. the chance of incurring Type I vs. Type II errors) determine people's reactions to possible cases of inductive risk. On the basis of these results, I discuss recent calls made by institutions to align research agendas to citizens' values, needs, and expectations.

Contents

Abstract	iv
1 Introduction	1
1 Values in science	6
1.1 Cognitive values	8
1.2 The desirability of cognitive values	10
1.3 Non-cognitive values	13
1.4 Non-cognitive values and theory appraisal: the descriptive, the epistemic, and the social	14
1.5 The epistemic-cognitive/non-cognitive distinction and terminology in the Science & Values debate	18
2 Methods	20
2.1 Case study: Evolutionary Psychology	21
2.2 Experimental philosophy	26
3 Summary and final remarks	28
2 What we (should) talk about when we talk about fruitfulness	29
1 Introduction	29
2 What is fruitfulness?	31
2.1 Fruitfulness as desirable development of programs	32
2.2 Developing the notion of fruitfulness	34
3 Fruitfulness at work: A case study	37
3.1 Research questions in Evolutionary Psychology	39
3.2 Heuristics in Evolutionary Psychology	42
3.3 Is Evolutionary Psychology fruitful?	44
4 Conclusion	45
3 The desirability of values in theory appraisal: A context-sensitive approach	47
1 Introduction	47
2 Explanatory power, external consistency, and fruitfulness of the Evolutionary Theory of Stalking	49
2.1 Stalking behaviours and sex differences	51

2.2	The Evolutionary Theory of Stalking	54
2.3	Explanatory power of the Evolutionary Theory of Stalking	57
2.4	External consistency and fruitfulness of the Evolutionary Theory of Stalking	65
3	Discussion: what the case study tells us about values	68
4	Conclusion	72
4	Feminist critiques of Evolutionary Psychology: The cognitive role of feminist values	73
1	Introduction	73
2	Human sexual behaviour and Evolutionary Psychology	74
2.1	Feminist critique of the evolutionary account of human mating	77
3	Similarities between cognitive and non-cognitive values	81
4	Conclusion	84
5	The psychology of inductive risk: An experimental study on the role of non-cognitive values in the assessment of cases of inductive risk	87
1	Introduction	87
2	Inductive Risk: The argument and psychological evidence	89
3	Experimental study on inductive risk	93
3.1	Experiment: Error type I	96
3.2	Experiment: Error type II	114
4	Discussion. The psychology of inductive risk	119
5	Conclusion	123
6	Conclusion	125
	Bibliography	130
	Acknowledgments	151

Chapter 1

Introduction

Female orgasm is “a sensory-motor reflex including clonic contractions (spasms) of the pelvic and genital muscle groups” involving pleasurable sensations (Lloyd 2005, 21). Female orgasm has puzzled evolutionary scientists for a long time and the reason is the following. While male orgasm is obviously connected to men’s reproductive success (i.e., success in spreading one's own genes), female orgasm seems to lack such a relation. Indeed, women can get pregnant without experiencing orgasm during intercourse, that is, female orgasm is not necessary for women’s reproductive success. So, why do women have the capacity to orgasm if this capacity does not obviously increase their reproductive success?

Another seemingly puzzling consideration is that (heterosexual) intercourse is not the ideal situation for women to orgasm. Among women, orgasm elicited by intercourse seems to be less frequent than orgasm caused by masturbation without vaginal insertion (Lloyd 2005, 37). Some studies also report that women involved in homosexual relationships experience orgasm during intercourse more often than women involved in heterosexual relations (86% vs. 65%, see Frederick et al. 2018). Moreover, no connection between female orgasm and women's fertility has been found, namely orgasmic capacity is not a cue of women's fertility (Lloyd 2005, 146).

Various theories have been formulated to explain the evolutionary origin of female orgasm. Although the available evidence does not suggest any connection between female orgasm and women’s reproductive success, the majority of these theories explains the trait as an adaptation, namely a trait evolved because it was able to enhance individuals’ reproductive success or survival. For instance, several theories understand female orgasm as facilitating the suction of sperm and semen in the uterus (e.g. Baker & Bellis 1993; Thornhill

et al. 1995). Such suction would facilitate fertilization, and then enhance women's reproductive success.

The theory formulated by the anthropologist Donald Symons constitutes an exception. Rather than explaining female orgasm as an adaptation, he suggested that female orgasm is a by-product of embryological development (Symons 1979). Female orgasm would then be one of those traits that sexes acquire in the embryological development, because strongly selected for one of the sexes. Females can orgasm, then, because orgasm is strongly selected for males for reproductive reasons, just like males have nipples because nipples are strongly selected in females for reproductive reasons.

By excluding a direct connection between reproductive success and female orgasm, Symons' theory can explain the available data on frequency and modalities of orgasm among women, and it can do it better than the adaptationist theories (namely the theories explaining the trait as an adaptation), for which the available evidence remains a puzzle (Lloyd 2005, 15). However, the theory of female orgasm as by-product has attracted little attention for a long time (Lloyd 2005; Lloyd 2015). Researchers have kept suggesting and endorsing theories explaining female orgasm as a trait enhancing reproductive success.

Various factors may explain why evolutionists have kept working on adaptationist theories of female orgasm despite the absence of a clear link between the trait and women's reproductive success. One factor may be the influence of the empirical and methodological assumptions known by the name of *adaptationism* (e.g. Godfrey-Smith 2001, 337). According to these assumptions, scientific theories that attempt to explain organisms' traits as adaptations should be preferred over theories explaining organisms' traits as other products of evolution (e.g. by-product); this preference would be grounded in the empirical idea that explanations of traits as adaptations are more likely to be true or with better chances to lead to the attainment of knowledge than explanations that do not posit that traits are adaptations.

Moreover, cultural and social assumptions may motivate the preference for the adaptationist theories of female orgasm over the non-adaptationist ones. For instance, Elizabeth Lloyd discusses how androcentric assumptions have influenced the research on the evolutionary origins of female orgasm. Research, Lloyd argues, has been guided by the idea that male sexuality is the fundamental model of sexuality: it provides the standard to which any form of sexuality should adhere. Any form of sexuality would then work in ways similar to that model and can be understood and explained by following that model (Lloyd 2005, 11).¹ Adaptationist theories would then be more attractive than the non-adaptationist ones

¹ Another possible cultural and social assumption influencing research on female orgasm is the idea that only heterosexual intercourse, and in particular the penis-vagina penetration, is sex. In this view, also called the "coital imperative", any other form of sexual activity is secondary or ancillary. Any trait associated to sexuality, such as female orgasm, should then be explained as having a function in heterosexual intercourse. See the feminist research on the coital imperative in McPhillis et al. 2001 and

because they would be able to make female orgasm fit the male model of orgasm (namely a trait directly connected to reproductive success).

In recent years, there has been an increasing interest among philosophers of science in studying how competing theories, such as the ones concerning female orgasm, can be evaluated and compared, and how social factors may influence such evaluations. Philosophers of science have focused on certain features of scientific theories, such as the power to explain the available observations and to unify bodies of evidence, arguing that these features are desirable for theories to have and relevant to the assessment and comparison of theories because indicative of the truth (McMullin 1983) or empirical adequacy of theories (van Fraassen 1980).

Various labels are used to refer to these features, such as theoretical virtues (McMullin 2009) and heuristics (Longino 2008). In this dissertation, I follow Laudan (2004) and use *cognitive values* to refer to these features. I refer to the social, political, religious, and economic factors that may influence scientific research as *non-cognitive values*. This dissertation is dedicated to exploring the importance of cognitive and non-cognitive values in the assessment and comparison of scientific theories, namely the stages of research in which scientists decide whether and which theories should be accepted or rejected.

While the philosophical literature has acknowledged the importance of cognitive values in these phases of research (see also Norton, unpublished), whether non-cognitive values can legitimately influence the assessment of theories has been hotly debated. Philosophers of science have traditionally argued that, if social, cultural or religious factors were to determine whether a theory should be accepted or rejected, science's ability to provide an objective account of the natural world and its epistemic authority in society would be compromised (Haack 1993). For instance, if the reason why some theories of the evolutionary origin of female orgasm are rejected is that they clash with some social beliefs and stereotypes about what female sexuality should be like, then it will be much more difficult that scientific research provides us with an objective understanding of women's sexuality.

However, the idea that the influence of non-cognitive values on theory appraisal is necessarily deleterious for science's epistemic authority has been widely criticized. A considerable amount of literature has been published, aimed at clarifying the legitimate, beneficial roles that non-cognitive values play in the assessment and comparison of scientific theories (e.g. Rudner 1953; Douglas 2000; Kourany 2010). Yet, several important questions concerning both cognitive and non-cognitive values remain to be addressed.

The importance of cognitive values for the assessment of scientific theories is widely recognized in the philosophical literature. Several lists of cognitive values have been suggested (e.g., Kuhn 1977; Longino 1995; Douglas 2009), but it remains unclear how

Jackson 1984.

exactly we should understand the nature and role of some of the values on those lists. Philosophers have clarified the importance and roles in theory appraisal of some values such as simplicity (Forster & Sober 2004), whereas little or no attention has been paid to other values, such as fruitfulness. Despite being frequently mentioned by both scientists and philosophers (e.g. Kuhn 1977; McMullin 1983; Douglas 2009), there is currently no careful explication of this value, and it remains unclear what it is that makes a theory fruitful, how we should make comparisons between the relative fruitfulness of competing theories, and when the relative fruitfulness of a theory is a good reason to accept the theory as true.

In the second chapter of this dissertation I will propose an explication of the value of fruitfulness and will sketch a strategy for assessing and comparing the relative degree of fruitfulness of competing scientific theories and research programs. Specifically, in the second chapter, I will address these questions:

What is the value of fruitfulness? How can we assess and compare the fruitfulness of alternative scientific theories and research programs?

I will then apply my explication and strategy to assess the fruitfulness of a research program, namely Evolutionary Psychology, which I will present in detail in section 2.1 of this introduction. I will talk about the fruitfulness of *research programs*, but my results can also be applied to scientific paradigms, theories, and models.

A further question concerns the *desirability* of certain values, namely the reasons why they constitute valuable features of scientific theories. As I have already mentioned, some philosophers of science argue that certain features of theories, such as fruitfulness, are desirable and relevant to the assessment of theories because they are indicative of the truth or empirical adequacy of theories. However, why is it so? In the third chapter, I will focus on the following question:

In virtue of what can a value provide us with reasons to believe that a given theory is true or empirically adequate?

The existing literature provides some answers to this question (e.g. Douglas 2009; Douglas 2013). However, the analysis of a case study will show that, in order to understand the ground for the desirability of cognitive values, at least some contextual factors (i.e. aspects of the context in which a value is used to assess theories) should be always considered. I will then argue that an accurate account of the legitimate roles that values can play in science should take into account that the desirability of values in science is always context-dependent.

At the beginning of this introduction, I described some of the research on female orgasm as an example of how social and cultural values can obstruct or limit an objective assessment of scientific theories. Social and cultural assumptions were among the factors

leading some evolutionists to prefer adaptationist theories of female orgasm, even though the available evidence did not show any apparent link between the trait and women's reproductive success. In the fourth and fifth chapter of this thesis, I will present two studies aimed at helping us understand the roles that non-cognitive values play in different scientific contexts.

In the fourth chapter, I will focus on theories of human mating within Evolutionary Psychology, and argue that non-cognitive values do not necessarily constitute a threat to the advancement of knowledge, but they can help reach an accurate assessment of theories. The chapter will revolve around the following question:

What are the epistemic benefits provided by non-cognitive values in the assessment of scientific theories?

To address this question, I will examine the influence of feminist values on Sexual Strategies Theory (Buss & Schmitt 1993). This case study shows that feminist values can foster the attainment of knowledge in several ways, such as by helping scientists to reveal when biases and assumptions are obstructing an accurate understanding of phenomena (cf., Fehr 2012).

In the fifth chapter, I will examine how non-cognitive values influence lay people's assessment of scientific hypotheses in cases of uncertainty. Uncertainty may be due to various reasons, such as the unavailability of unambiguous results or disagreement among scientists on the reliability of scientific methodologies. In these cases, scientists may wrongly assess scientific hypotheses (e.g. accepting a hypothesis that should be rejected). Mistakes can have consequences that can be morally or economically undesirable and non-cognitive values, some philosophers argue, provide the standards to evaluate and compare these possible consequences (e.g. Rudner 1953; Douglas 2000).

However, little is known on how - specifically - non-cognitive values influence this evaluation. The following question will be the focus of the chapter:

What is the influence of non-cognitive values on people's scientific reasoning in cases of uncertainty?

To answer the question, I will present and discuss the results of an empirical study I carried out with the aim of clarifying the role of political values and personal features (e.g. sex and race) in reasoning about situations of uncertainty in science.

The central questions of this dissertation will then be addressed by employing two main methods, namely the analysis of a case study (Evolutionary Psychology) and the experimental method (a survey). I will present these methods in the final pages of this introduction. Before doing so, in the next sections I will discuss the Science & Values debate, which constitutes the basis of the arguments developed in the next chapters.

1. Values in science

Social and cultural assumptions about female and male sexuality were among the factors that led some scientists to prefer theories explaining female orgasm as an adaptation in spite of the little evidence supporting this view. Philosophers of science have traditionally argued that, in order to avoid cases like this one, the assessment and comparison of theories should be a value-free enterprise, in particular, it should be free from the influence of non-cognitive values (e.g. Haack 1993). The idea that science should be kept free from the influence of non-cognitive values is commonly referred to as the *value-free ideal of science* (e.g. Reiss & Sprenger 2017).²

The ideal concerns the roles of non-cognitive values in some specific stages of research, namely the assessment and comparison of scientific theories. This means that the influence of non-cognitive values on other stages of research, such as the choice of research questions and methodologies, is seen as indisputable and unproblematic. For instance, cultural and social values may inspire and motivate scientists to focus on questions concerning the adaptive origin of female orgasm. Ethical values may be relevant in methodological decisions. Because of ethical concerns, scientists may decide to avoid some methods to collect data on female orgasm if these can threaten the dignity and health of the women involved in the studies. The influence of non-cognitive values on these stages of research is not seen as a threat to the epistemic authority of science.³

I have presented the value-free ideal as an ideal limiting and disciplining the influence of *non-cognitive values* on science. However, there are various versions of the value-free ideal and one of them concerns both cognitive and non-cognitive values.

Richard Jeffrey (1956) has defended this version of the value-free ideal and argued that, when comparing rival hypotheses, scientists make no value judgments. When comparing rival hypotheses, scientists just assign probabilities to hypotheses on the basis of the available evidence (Jeffrey 1956, 237). Isaac Levi (1960) has argued in favour of this version of the value-free ideal by pointing out that a scientist, when committing herself to scientific research, also commits herself to certain standards of inference, which guide scientific reasoning (e.g. establishing levels of statistical significance) and have been decided by the scientific community. The scientist is “obligated to accept the validity of certain types

² The literature provides various versions of the value-free ideal in which non-cognitive values have different levels of importance (see for instance Dorato 2004). For an articulated explication of the value-free ideal see Lacey 1999, who sees the value-free ideal as composed by three main claims concerning the impartiality, neutrality, and autonomy of science from non-cognitive values.

³ The role of non-cognitive values in these stages has recently started to attract more attention. Kevin Elliott and Daniel McKaughan have argued that, by influencing the choices of projects and questions, the development of theories, and the collection of data, non-cognitive values end up significantly affecting the assessment and comparison of theories (Elliott & McKaughan 2009).

of inference and to deny the validity of others” because of these standards of inference (Levi 1960, 356). According to Levi, this means that when accepting or rejecting hypotheses, scientists do not make any value judgments. Rather, they simply apply standards of inference on which the scientific community has agreed.

Currently, few philosophers support this version of the value-free ideal. Several studies in history, sociology, and psychology of science have emphasized how value judgments are intrinsically embedded in scientific practice (e.g. Ruse 1999; see also Norton, unpublished) and it is then descriptively inaccurate to claim they do not have any role in scientific reasoning.

The current debate focuses on the version of the value-free ideal limiting the roles of non-cognitive values. In this view, scientists make value judgments. Specifically, they make value judgments involving cognitive values when assessing and comparing rival scientific theories. However, non-cognitive values are excluded from having any legitimate, beneficial role in the assessment of scientific theories (e.g. McMullin 1983; Haack 1993) since they are *irrelevant* to the assessment of theories. Non-cognitive values, advocates of the value-free ideal argue, are unrelated to the truth or empirical adequacy of theories, and this means that they cannot provide any relevant information to assess the truth or empirical adequacy of theories. When they influence the assessment of theories, non-cognitive values would then have *detrimental* effects on science, misleading or undermining the reliability of scientific reasoning. Susan Haack (1993) supports this version of the value-free ideal and emphasizes how the influence of non-cognitive values on the assessment of theories may have epistemically and socially undesirable consequences like dogmatism. Impeding the influence of non-cognitive values on theory choice, then, is “a matter, epistemologically, of the integrity of inquiry and, politically, of freedom of thought” (Haack 1993, 38).

The value-free ideal is an attractive ideal. It provides us with a clear image of science and it specifies how science should be like if it is to provide us with an objective, epistemically authoritative understanding of the natural world. Moreover, it protects scientific practice from the influence of economic, social, and political interests, that may undermine human beings' freedom of thought.

However, as the arguments developed in this dissertation will show, the value-free ideal of science fails to account for some important, beneficial roles that values play in science. Two dimensions of my analysis should be clear. First, my analysis makes a *descriptive* claim. The value-free ideal of science provides an inaccurate portrayal of scientific reasoning, as Chapters 4 and 5 will demonstrate. Second, in my analysis I make a *normative* claim: the value-free ideal does not constitute a desirable ideal. Both cognitive and non-cognitive values constitute essential and often beneficial aspects of science, and eliminating them in all cases means removing important factors operating in scientific research.

In the next section, I will present some of the existing arguments defending the

legitimate, beneficial roles that values play in scientific research and in particular in the assessment of scientific theories.

1.1 Cognitive values

Some philosophers of science argue that certain features of scientific theories, such as simplicity and external consistency, are valuable because it is epistemically desirable, for scientific theories, to satisfy them (McMullin 1983, 5). Specifically, they are valuable because indicative of the truth or empirical adequacy of scientific theories.

Philosophers argue that these features, or cognitive values, play a fundamental role in science.⁴ For instance, Thomas Kuhn states that “abandoning them would be abandoning science altogether” (Kuhn 1993, 331-2) and Larry Laudan that “we cannot conceive of a functioning science without them” (Laudan 2004, 19). Cognitive values would then play a fundamental role in the assessment of scientific theories, constituting “the shared basis for theory choice” (Kuhn 1977, 357). In particular, they would help scientists in cases of underdetermination, in which scientists have competing theories that are empirically equivalent to each other (i.e. they are all compatible with the available evidence) but are incompatible with each other, namely they account for the data in hand in conflicting ways. In these cases, cognitive values provide reasons to prefer one theory over its alternatives.

Kuhn explains this point in his paper “Objectivity, Value Judgment, and Theory Choice”. In this paper, he draws one of the first lists of cognitive values (Kuhn 1977, 322), which includes accuracy, consistency, scope, simplicity, and fruitfulness. He argues that, when assessing alternative theories, accuracy is usually taken as the most fundamental value. This is because the value of accuracy concerns the agreement between the theory and “the results of existing experiments and observations” (Kuhn 1977, 321), and this is taken as an essential virtue of scientific theories. However, he argues that considering accuracy may be insufficient to understand which of two rival scientific theories should be preferred. Other factors may be considered in order to decide which theory should be accepted. To illustrate this point, Kuhn uses the example of Copernicus' and Ptolemy's models of the universe:

“Copernicus's system [...] was not more accurate than Ptolemy's until drastically revised by Kepler more than sixty years after Copernicus's death. If Kepler or someone else had not found other reasons to choose heliocentric astronomy, those improvements in accuracy would never have been made, and Copernicus's work might have been forgotten. More typically, of course, accuracy does permit

⁴ Various labels have been used to refer to these features. For instance, they are called ‘theoretical virtues’ in van Fraassen 1980 and McMullin 2008, while Longino 2008 uses the term ‘heuristics’. Kuhn (1977) initiated the use of calling these features *values*, and this is now the common practice in the literature (e.g. Rolin 1998; Douglas 2000; Douglas 2009; Steel 2010; Rooney 2017; see also Norton, unpublished).

discriminations, but not the sort that leads regularly to unequivocal choice.” (Kuhn 1977, 323)

The “other reasons” were provided by some cognitive values. For instance, Kuhn points out that simplicity favoured Copernicus' over Ptolemy's model. Specifically, Copernicus' model required a simpler mathematical apparatus, since it assigned just one orbit for each planet, compared to Ptolemy's one, which assigned two (Kuhn 1977, 324). Kuhn argues that these considerations about simplicity, among others, played a fundamental role in the comparison of the two models. Indeed, he claims that a long look at the history of science reveals these considerations were “vitally important to the choices made by both Kepler and Galileo and thus essential to the ultimate triumph of Copernicanism” (Kuhn 1977, 324; see also Norton, unpublished).

Philosophers have suggested various lists of cognitive values. In table 1, I provide an overview of some of these lists. It is interesting to notice that, despite having different ideas on the aims of science, philosophers agree on the importance of the very same values in the assessment of scientific theories. Indeed, most of the lists significantly overlap.

For instance, McMullin and Kuhn suggest almost identical lists, but while McMullin endorses scientific realism (roughly speaking, he believes science's aim is providing true theories and science is able to do so), Kuhn does not. This difference is mirrored in what these values are indicative of in McMullin's and Kuhn's accounts. While for McMullin features as fruitfulness and consistency are truth-indicative features of theories, i.e., informative over the truth value of theories, Kuhn argues these values indicate the ability of a theory to solve puzzles.

Kuhn 1977	Accuracy, consistency, scope, simplicity, fruitfulness.
McMullin 1983	Predictive accuracy, internal consistency, external consistency, unifying power, simplicity, fruitfulness.
Longino 1995	Empirical adequacy, novelty, ontological heterogeneity, complexity and mutuality of interactions, application to human needs, diffusion of power.
Douglas 2009	Predictive competence, internal consistency, simplicity, explanatory power, scope, external consistency, predictive precision, fruitfulness.

Table 1. Lists of values.

The list suggested by Helen Longino, instead, considerably differs from the others (Longino 1995; Longino 2008). In her list, Longino includes values clashing with the traditional ones (e.g. novelty vs. external consistency, ontological heterogeneity vs. simplicity). The only exception is empirical adequacy, which figures in the traditional as well as the alternative list. While the traditional lists include external consistency, which recommends preferring theory consistent with the background theories over the alternatives that are in contrast with them, Longino's list includes the value of novelty, which, as Longino understands it, pushes towards contrast. According to Longino, the theories proposing novel frameworks of understanding that contrast with the background theories should be preferred over the ones consistent with the background theories. Longino includes these alternative values in her list because these values, in some circumstances, may perform better in promoting the attainment of knowledge than the traditional ones (Longino 1995, 385).⁵ Feminist research in primatology is a case in point. Feminist primatologists have preferred to formulate and work with theories proposing a novel framework of understanding to the ones consistent with the background theories. This, Longino argues, has led to a better understanding of primates' behaviours. Specifically, by contrasting the traditional view in which the male primate was seen as the representative entity of the members of the primates' community, feminist researchers were able to achieve a better understanding of the actual differences among the members of primates' communities (Longino 2008, 70).

Longino's analysis reveals how different contexts may make different values desirable and able to guide scientists towards the attainment of knowledge. I take up and develop this insight throughout my thesis.

All in all, despite some disagreement on what a list of values should include, philosophers of science agree that certain features of theories play an important role in the assessment of scientific theories. In the next subsection, I will present some of the argumentative strategies used by philosophers to clarify the desirability of cognitive values in science.

1.2 The desirability of cognitive values

One of the first strategies philosophers of science have used to clarify the desirability of cognitive values and their importance for theory appraisal is a historical one. McMullin adopts this strategy and claims that history of science shows that "the assessment of theories involves value-judgments in an essential way" (McMullin 1983, 14):

⁵ Paul Feyerabend makes similar remarks when claiming that the traditional values may impede the production of knowledge and that allowing alternative criteria to guide theory choice may enhance scientific progress (Feyerabend 1975, 15).

“The characteristic values guiding theory-choice are firmly rooted in the complex learning experience which is the history of science; this is their primary justification and it is an adequate one.” (McMullin 1983, 21)

According to McMullin, history of science shows that scientists do use values for assessing and comparing scientific theories and that using these values to assess and compare theories have enhanced the attainment of scientific knowledge. Certain values are desirable, then, because they are truth indicative: they have been successful in helping scientists to choose true theories and reject the false alternatives. The fact that they have been “epistemically successful” in the history of science would demonstrate their status of epistemically desirable features of scientific theories (see also Rolin 1998).

The historical argument challenges the value-free ideal presented in the previous section, by showing that scientists make value judgments and that these value judgments are not epistemically deleterious to science. Several studies have discussed the importance of values in specific episodes of the history of science. For instance, Michael Ruse has analysed the development of evolutionary theory in the history of science and argued that “epistemic norms play a major role in the structure of evolutionary theorizing” (Ruse 1999, 237). Some studies report the explicit appeal of scientists to specific values (e.g. Zahar 1973 on the importance Einstein ascribed to simplicity and coherence in the assessment of theories), whereas others, when scientists’ appeal to values is not so explicit, engage in an interpretative activity (see for instance Nolan 1997 on the role of ontological parsimony in the work of Wolfgang Pauli and Enrico Fermi).⁶

Besides appealing to the history of science to argue for the desirability of cognitive values, some philosophers have tried to clarify the specific reasons for the desirability of each value (e.g. Douglas 2009; Douglas 2013). Two main reasons have been suggested:

- (a) A value is desirable because it is truth indicative.
- (b) A value is desirable because of pragmatic reasons.

To refer to these two kinds of values, from now on, I will use the notation suggested in Laudan (2004). I will call *epistemic* values those values that are desirable because truth indicative and *cognitive* value those values that are desirable because of pragmatic reasons. When not specified otherwise, I will use ‘cognitive values’ to refer to both epistemic and cognitive values (e.g. the discussion in Chapter 4).

Being truth indicative means being informative over the truth value of theories. If a

⁶ The interpretative activity has also produced some disagreement among philosophers. In some cases, philosophers disagree on which specific value has played a central role in a certain episode from the history of science. See Margaret Morrison (2000) and Samuel Schindler (2014) as an example of disagreement over the importance of the value of unification in the development and success of the Glashow-Weinberg-Salam model of electroweak interactions.

theory satisfies this value, such a theory is probably true or approximately true. Truth-indicative values, then, are desirable because they help scientists detect true or approximately true theories among competing theories. Predictive accuracy, namely the degree of agreement between the predictions scientists draw from a theory and what scientists actually observe in the world, is usually listed among the truth-indicative values. For McMullin, all the values included in his list are truth-indicative values (see table 1). In his view, values such as fruitfulness and scope “promote the truth-like character of science” and “we have reason to believe will, if pursued, help toward the attainment of [...] knowledge” (McMullin 1983, 18). As mentioned in the previous section, McMullin endorses a realist view of science. However, it should be noted that no commitment to scientific realism is necessarily involved here. As Daniel Steel explains, epistemic values may indicate that the theory is able to provide true predictions about observables (Steel 2010, 18), which is coherent with certain kinds of scientific antirealism, such as constructive empiricism (see van Fraassen 1980).

Pragmatic reasons constitute another source for the desirability values in science. There are at least two (partially overlapping) categories of pragmatic reasons grounding the desirability of a value. First, values may be desirable because they provide some *scientific* pragmatic benefits. For instance, when talking about the desirability of simplicity, Heather Douglas argues that “simpler claims are easier to follow through to their implications” (Douglas 2013, 800). The desirability of simplicity is then attributable to the pragmatic benefit of making the scientific work easier. This means that certain values may be desirable since they “help one think through the evidential and inferential aspects of one's theories and data” (Douglas 2009, 93).

Second, the desirability of a value can be explained in terms of *personal or social* pragmatic benefits. A value may be desirable because it enhances the reputation of scientists or it facilitates meeting some societal needs. For instance, Kuhn explains the desirability of fruitfulness by referring to the reputation of scientists. He construes fruitfulness as the power to produce new research findings and claims that one of the reasons of its desirability is that it enables scientists to achieve rewards and popularity in a certain field:

“a scientist choosing between two theories ordinarily knows that his decision will have a bearing on his subsequent research career. Of course he is especially attracted by a theory that promises the concrete successes for which scientists are ordinarily rewarded.” (Kuhn 1977, 322)

Moreover, Longino has analysed the desirability in terms of the social benefits provided by the alternative values she suggests (see table 1 in the previous section). Values as novelty and ontological heterogeneity were found desirable by researchers because able to serve feminist goals (Longino 2008, 77). That is, they contributed to dismantle stereotypes about gender,

and to detect sexist bias in scientific theories (Longino 1995, 391).⁷

The two kinds of reasons explaining the desirability of values are not mutually exclusive. A value may be desirable because it is truth indicative *and* because of some pragmatic reason. For instance, external consistency (i.e. the absence of contradictions between the theory under scrutiny and other scientific theories) may be desirable because truth indicative. If we are confident that the other theories are true, then the consistency of our theory with other theories can be an indicator that the theory is probably true or at least not obviously false. However, the desirability of external consistency may involve pragmatic reasons as well, such as economic reasons. The value of external consistency may be desirable because theories satisfying this value would make it easy for scientists to receive funds for their research and to reach support and a broad consensus in the scientific community.

The third chapter of this dissertation will be dedicated to exploring what it is that makes values desirable features of scientific theories. Specifically, I will analyse the role of contextual factors in making some values desirable for epistemic reasons. Let's now turn to the literature on the importance of non-cognitive values in science.

1.3 Non-Cognitive Values

Non-cognitive values constitute a heterogeneous group.⁸ In this group, philosophers include all the “extra-scientific” values, such as moral, social, religious, political, aesthetic, and economic values. Claiming that these values are *extra-scientific* does not mean these values are (all) unrelated to science.⁹ These values are seen as extra-scientific since they are not related to the *scientific goals* of research, such as providing an objective account of the world. However, the distinction among kinds of values is far from being a trivial matter. It is not clear whether it is possible to provide a clear demarcation between epistemic, cognitive, and non-cognitive values (Rooney 1992; Rooney 2017; Steel 2010). This issue will be analysed in section 1.5 of this chapter.

As discussed in the first pages on this introduction, advocates of the value-free ideal argue that the influence of non-cognitive values on the assessment and comparison of scientific theories has always deleterious effects. In particular, it would have epistemically undesirable consequences (compromising science's ability to provide an objective understanding of the world) as well as socially undesirable consequences (threatening freedom of thought) (Haack 1993; Dorato 2004).

⁷ Longino also discusses how the influence of extra-scientific interests on the desirability of cognitive values may compromise scientific practice (Longino 1995, 393).

⁸ As for cognitive values, philosophers have used different labels to refer to this kind of values, such as non-epistemic (e.g. Rooney 1992) and contextual values (Longino 1990).

⁹ Some non-cognitive values can even be developed in academic contexts. For instance, aesthetic and political values can be the product of research in aesthetics, political studies, and ethics.

Feminist philosophers have been among the most vocal scholars highlighting the deleterious effects of non-cognitive values on science. Their criticism has focused on the influence of sexist, androcentric, racist, and ethnocentric values on science (e.g. Anderson 2004). For instance, some feminist philosophers have pointed out how research in biology and primatology has been guided by sexist and androcentric values (e.g. Biology and Gender Group Study 1988; see also Chapter 4 of this dissertation) and how these values have obstructed efforts to understand various phenomena (e.g. fertilization, roles of female primates in primate communities). However, some feminist researchers have also pointed out that the research informed by feminist values has performed better than the one influenced by androcentric assumptions in terms of scientific standards, such as empirical accuracy (e.g. Richardson 2008). The research conducted by scientists explicitly committed to feminist values, feminists argues, has been able to provide us with a better understanding of biological phenomena.

However, feminist values are as “extra-scientific” as sexist and androcentric values are (Antony 1993). They all belong to the group of non-cognitive values that are supposed to be unrelated to the scientific goals of research. To explain this asymmetry in the effects of different kinds of values on science, i.e., to explain why some non-cognitive values are beneficial to science whereas others undermine its epistemic authority, philosophers of science have suggested various strategies (e.g. Hicks 2014; Psillos 2015; Goldenberg 2015; Brigandt 2015). The fourth chapter will address this issue by analysing the epistemically beneficial roles that feminist values have played in the development and critique of theories of human mating in Evolutionary Psychology.

A large and growing body of literature has defended the legitimate, beneficial roles that non-cognitive values can play in the assessment of theories (e.g. Brown 2013). Moreover, some philosophers have suggested ideals that may replace the value-free ideal of science, such as the social management ideal of science (Longino 2002) and the ideal of socially responsible science (Kourany 2010). It is beyond the scope of this dissertation to present and review all the arguments and ideals suggested. However, I will present three main claims that underlie these arguments, namely a descriptive claim, an epistemic claim, and a social claim. These claims will be relevant to the discussion of non-cognitive values I will develop in the fourth and fifth chapters of this dissertation.

1.4 Non-cognitive values and theory appraisal: the descriptive, the epistemic, and the social

The first claim made by the arguments defending non-cognitive values is a descriptive one: a long look at scientific practice reveals that non-cognitive values constitute an important component of scientific reasoning and the assessment and comparison of scientific theories. Philosophers argue there should be little doubt that this is so. Research in philosophy, sociology, and history of science has shown that “the work of even the greatest scientists [...]

has been shaped by social values” (Kourany 2010, 55).

Value-free ideal advocates do not reject this descriptive claim. Both advocates and critics of the value-free ideal recognize that history of science shows that non-cognitive values exert an influence on scientists’ decisions of accepting or rejecting theories and hypotheses. However, they disagree on what history of science teaches us about the influence of non-cognitive values on science. Advocates of the value-free ideal argue that the history of science reveals that non-cognitive values obstruct the epistemic aims of scientific research by leading to a partial and inaccurate understanding of phenomena. In contrast, according to the critics of the value-free ideal, instead, the history of science shows that non-cognitive values can have a positive influence on science, by providing epistemic and/or social benefits. These points will be explored in the next paragraphs, when I will present the epistemic claim and the social claim. Before doing that, let’s have a closer look at one of the ways in which non-cognitive values influence scientific research, namely the role that non-cognitive values play in evidential reasoning, as presented in Longino (1990).

Longino argues that non-cognitive values (or contextual values, as she calls them, Longino 1990, 4) are involved in the way evidential reasoning works (Longino 1990, 43). The reason is that there is no direct relation between states of affairs and hypotheses: the same data on certain phenomena can support alternative hypotheses explaining those phenomena. States of affairs “do not carry labels indicating that for which they are evidence or for which they can be taken as evidence” (Longino 1990, 40). Longino argues that the background assumptions endorsed by scientists, such as metaphysical beliefs and social, ethical, and political values, bridge the gap between data and hypotheses. Whether the available data are taken as evidence of a specific hypothesis is dependent on the background assumptions endorsed in the context in which data are considered (Longino 1990, 43). For instance, whether data on modalities of orgasm among women are taken as evidence for an adaptationist theory of female orgasm may depend on the beliefs and assumptions endorsed by scientists on how female sexuality should be like.

According to Longino, background assumptions are *required* “to demonstrate the evidential import of a set of data to a hypothesis” (Longino 1990, 59). Evidential reasoning, then, reveals that background assumptions play a fundamental role in science: without them no relation between evidence and hypotheses can be determined. However, although non-cognitive values are involved in how evidential reasoning works, it may be argued that their influence is detrimental to the epistemic authority of science. In other words, it may be argued that, granted, non-cognitive values are involved in evidential reasoning, but we should try to limit their influence. This leads us to introduce the epistemic claim.

The epistemic claim is that non-cognitive values *should not* always be excluded from having an influence on the assessment of theories because in certain circumstances they may provide epistemic benefits. In other words, non-cognitive values are not necessarily an obstacle to the attainment of knowledge. Rather, they can foster it.

For instance, Longino argues that the fact that non-cognitive values influence evidential reasoning does not mean that their influence should be always minimized. More precisely, the epistemic authority of science is not preserved by eliminating non-cognitive values from science, but through a social process of criticism in which the influence of non-cognitive values on science is socially disciplined. According to Longino, scientific knowledge is “produced collectively through the clashing and meshing of a variety of points of view” (Longino 1990, 69). Intersubjective criticism among scientists and communities is what makes achieving an accurate understanding of the empirical world possible (Longino 1990, 71). Indeed, according to Longino the background assumptions endorsed by scientists can be articulated, clarified, and - if necessary - criticized. Moreover, Longino argues that scientific communities should involve individuals and groups with heterogeneous assumptions, beliefs, and values, which would foster an epistemically successful intersubjective criticism.

I will discuss the epistemic benefits of non-cognitive values in the fourth chapter, where I will analyse the influence of feminist values on research on human mating. Let’s now consider the last claim.

The third claim is a social claim: the assessment and comparison of theories has a socially and ethically relevant dimension in which non-cognitive values can play a legitimate role. Specifically, some philosophers argue that there are decisions in science for which the scientific method cannot provide an answer (Hempel 1965, 86). Providing an answer to these questions should involve non-cognitive values.

This is the case of the assessment of theories done in the face of uncertainty. Uncertainty may be due to the lack of relevant evidence, mixed results or disagreement over scientific methodologies. Because of uncertainty, scientists may make mistakes. For instance, they may mistakenly accept a false hypothesis or reject a true hypothesis. Philosophers of science call the chance of making mistakes when assessing hypotheses in the face of uncertainty *inductive risk*. There is the inductive risk that future evidence will reveal that the present assessment is mistaken (Hempel 1965).

Mistaken decisions may have various consequences, which may be undesirable for social, economic, and moral reasons. To illustrate, consider the female orgasm debate. Accepting the hypothesis that female orgasm is an adaptation, when this is actually false, may have socially relevant consequences. For example, low women’s orgasmic capacity in intercourse may be seen as dysfunctional, which will have consequences for “how women’s sexuality is socially and personally perceived” (Lloyd 2005, 19).

The level of confidence requested for accepting (or rejecting) a scientific hypothesis “depend[s] on how serious a mistake would be” (Rudner 1953, 3) and the seriousness of the mistake and its consequences depends on social, economic, and moral standards. Deciding how strong the evidence has to be in order to accept (or reject) a hypothesis involves then weighing the possible social, economic, and moral consequences of a mistaken assessment.

The idea, then, is that when moral or social consequences are involved, non-cognitive values are required in scientists' and policy makers' decisions, they are needed for good scientific reasoning (Douglas 2000, 565). However, non-cognitive values do not determine whether a hypothesis is true. Advocates of the argument from inductive risk argue that values and evidence play different roles in the assessment of scientific hypotheses: non-cognitive values do not constitute evidence and their importance in the assessment of hypotheses is secondary to the importance of evidence (Douglas 2009, 87). The role of non-cognitive values is limited to assisting scientists when considering and weighing the possible consequences of errors, helping to find the most acceptable course of action (for their moral, social or economic standards) in cases of uncertainty.

The argument from inductive risk does not concern the assessment of hypotheses only. Inductive risk permeates various stages of scientific research. Scientists may make mistakes when choosing methodologies, gathering and characterizing data, and interpreting evidence, and these decisions may have consequences that may be more or less desirable for our moral, social or economic standards (Douglas 2000, 565). Thus, non-cognitive values play an important role at all these stages.

Inductive risk and the role of non-cognitive values in reasoning involving uncertainty will be explored in the fifth chapter of this dissertation. However, it should be noted that there is disagreement on whether the argument from inductive risk can actually show that non-cognitive values play a legitimate, beneficial role in science. For instance, McMullin claims that considering the ethical implications of the assessment of hypotheses does not belong to scientists' job (McMullin 1983). When scientists accept a hypothesis, he argues, they do so because it is "the best-supported of the alternatives available or [...] offering the most fruitful research program" (McMullin 1983, 8), i.e., they evaluate the likelihood or fruitfulness of the hypothesis. Considering the possible consequences of an error does not belong to the scientific method and scientists are not "called upon to make [ethical] value judgments" (McMullin 1983, 8; see also Jeffrey 1956, Levi 1960).

In a similar vein, John Norton has argued, paraphrasing the title of Richard Rudner's paper on inductive risk "The scientists qua scientists make value judgments", that scientists qua scientists do not make ethical judgments. Rather, scientists *qua members of society* make value judgments. However, these criticisms seem to ignore the moral responsibilities of scientists. Are scientists allowed to ignore the possible ethical and social consequences of their decisions and ignore the fact that science is done in a social context?

The argument from inductive risk will be explored in the fifth chapter, in which I will present an experimental study examining the importance of non-cognitive values in the assessment of cases of inductive risk. Before presenting the methods used in this dissertation, a final remark on the distinction of values and terminology I use.

1.5 The epistemic-cognitive/non-cognitive distinction and terminology in the Science & Values debate

The epistemic/cognitive-non-cognitive distinction is not widely accepted among philosophers of science. It has been criticized as vague and problematic. Indeed, it is not clear what should count as an epistemic/cognitive value and what should be seen as a non-cognitive value. For instance, what should be included in the group of epistemic and cognitive values (i.e. the values related to the scientific goals of science or enhancing the attainment of knowledge)?

Phyllis Rooney has argued that philosophers significantly disagree on this issue, suggesting lists including different values. She mentions Kuhn's and McMullin's lists of values (see table 1) and the list of values suggested in Longino 1990 (which includes truth, accuracy, simplicity, predictability, and breadth; Longino 1990, 4). It is interesting to notice that while Rooney argues that these lists differ considerably, I have claimed that they significantly overlap (with the exception of the list suggested in Longino 1995; see section 2 of this introduction). The values of accuracy in Kuhn 1977, predictive accuracy in McMullin 1982 and accuracy in Longino 1990 seem to refer to similar features of scientific theories. Moreover, all these lists include values concerning the simplicity of theories and the breath of theories (e.g. scope and unifying power in Kuhn 1977 and McMullin 1983). So, it is not obviously true that, as Rooney suggests, there is disagreement about what values are possibly relevant in science.

Moreover, Rooney criticizes the cognitive/non-cognitive distinction by pointing out that some values commonly understood as non-cognitive values, in certain contexts, may promote the attainment of knowledge (as it was also pointed out in the epistemic claim introduced in the previous section). For instance, she mentions the epistemically beneficial influence of theological beliefs on Newton's work in celestial mechanics (Rooney 1992, 16; see also McMullin 1983, 19-20).¹⁰ If religious beliefs and values can help us attain knowledge of the natural world, why should they be excluded from the group of epistemic and cognitive values?

Rooney concludes there is no clear borderline between cognitive and non-cognitive values and the distinction should be abandoned. The distinction is not useful to understand how science works in the social context (Rooney 1992, 13) and to account for the actual development of science (Rooney 1992, 17). She then suggests to replace the distinction with a continuum scale that could account for the different degrees in which the values commonly understood as epistemic, cognitive, and non-cognitive can enhance the attainment of knowledge. The epistemic importance of values, Rooney argues, may be subjected to renegotiation depending on the specific aims of the context in which the values are employed

¹⁰ A further example of values traditionally construed as non-cognitive but having an epistemic role are aesthetic values (see Ivanova 2017).

(Rooney 1992, 21; see also Rooney 2017).

The cognitive role that some non-cognitive values may have in certain contexts will be addressed in the fourth chapter of this dissertation. Before going further with presenting the methods used to explore the questions addressed in this dissertation, a final remark on the terminology used in the Science & Values literature.

I have adopted the term ‘value’ to refer to the desirable features of scientific theories. As mentioned, different terms have been adopted, but using the term ‘value’ is the common practice in the philosophical literature (e.g. Kuhn 1977; McMullin 1983; Douglas 2013). However, this terminological choice has been criticized.

For instance, Norton has argued that the term “value” does not capture what features like simplicity and fruitfulness really are (Norton, unpublished). He argues that, since values are “commonly understood to be things that we esteem in their own right” (unpublished, 2), values should generally be understood as *ends*. But, according to Norton, features such as simplicity and fruitfulness are not scientific ends, but *means* to pursue some scientific goal. Specifically, they are means to achieve the end of getting closer to truth (Norton unpublished, 14). Calling them values is a mistaken terminological choice that may lead to misunderstandings of what these features really are and of their actual role in scientific practice.

When Norton claims that values are “commonly understood to be things that we esteem in their own right,” he does not specify whether he refers to folks’ ordinary understanding of values or to some technical meaning employed by value theorists. Depending on which context he refers to, his criticism may be justified or not. Indeed, it is not true that values are understood as ends if we consider Ethics. Ethicists talk about values in terms of intrinsic and extrinsic values. While something has intrinsic value if it is valuable in its own right (the notion of values Norton seems to refer to), having extrinsic value means being valuable because of external factors. Specifically, what has extrinsic value is valuable because of its relation to something else which is valuable. For instance, hedonists think that while the only thing with intrinsic value is pleasure, the things leading to the attainment of pleasure have extrinsic value. The value of the things leading to the attainment of pleasure is then subordinate and parasitic on the value of pleasure (Harman 2000, 137).

Although simplicity and fruitfulness may be not valuable *per se*, as Norton argues, and should be construed as a means to reach some important goals (e.g. the attainment of knowledge), the term “value” can still be used to refer to them. Indeed, features as simplicity and fruitfulness may be seen as having extrinsic value. In particular, these features may be seen as having a specific kind of extrinsic value, namely *instrumental* value, which means being valuable because helping to reach other goods we find valuable (Driver 2006, 8). Simplicity and fruitfulness, then, would be valuable because instrumentally helpful to reach other values, such as accuracy.

Critics may argue that the term ‘value’, when used in the everyday context, does not

include the notion of instrumental value, and that, when we talk about values, we talk about things having value in their own right. Critics may conclude, then, that the term value should not be used to talk about features of scientific theories relevant to the assessment of theories. It is beyond the scope of this dissertation understanding how the term value is used in everyday context. However, it is not clear why we should take the everyday context and natural language as the context relevant to establish the appropriate terminology in the Science & Values debate. Moreover, the terminology and concepts developed in Ethics can be useful to explore the nature and importance of values in science. Indeed, the notions of intrinsic, extrinsic, and instrumental values may constitute useful tools to understand the relations between cognitive values and their roles in theory choice (see for example the analysis developed in Steel 2010).

In the final pages of this introduction, I will present the methods used in my analysis of the importance of epistemic, cognitive, and non-cognitive values in science.

2. Methods

Appealing to case studies, namely episodes of the history of science or current scientific practice, is the main method used by philosophers to discuss the roles of values in science. For instance, Elisabeth Anderson has used a case study from research on divorce to examine how non-cognitive values may guide scientific research, e.g. by determining the way the object of inquiry is conceived and what conclusions are drawn from the evidence collected (Anderson 2004, 13-18). Daniel Nolan has analysed the role of quantitative parsimony in the work of Wolfgang Pauli and Enrico Fermi that led to the discovery of the neutrino (Nolan 1997), and argued that the preference for quantitatively parsimonious hypotheses have helped scientists to “move closer to the truth” (Nolan 1997, 342). Using this approach, philosophers have been able to reveal that (epistemic/cognitive/non-cognitive) values influence science and to point out that their influence has been (epistemically or socially) beneficial or detrimental to science.

In this dissertation, I make use of a case study to explore the roles of epistemic, cognitive, and non-cognitive values in science, namely the research on the evolutionary origins of human psychology. Various approaches to the study of the evolution of human psychology have been developed. These approaches may differ in the methodologies used and the theoretical tenets accepted (see Scher and Raucher 2003 for an overview of the different approaches). The case study presented in this dissertation focuses on a specific line of research. Various labels have been used to refer to this school, such as “inclusive fitness psychology” (Caporael 2001) and “narrow evolutionary psychology” (Scher & Raucher 2003). I will follow the notation used in Buller (2005), namely *Evolutionary Psychology* with capital letters. Some of its most prominent researchers are Leda Cosmides and John Tooby,

who work at and have founded the Center for Evolutionary Psychology at the University of California, and David Buss, who leads and has founded the Buss Lab at the University of Texas.

I have used this case study for answering three of the questions explored in this dissertation. In the second chapter, Evolutionary Psychology is used to understand how fruitfulness can be explicated and evaluated. In the third chapter, I analyse a theory developed in this research program, namely the Evolutionary Theory of Stalking (Duntley & Buss 2012), to understand the reasons of the desirability of some values (explanatory power, external consistency, and fruitfulness) and whether contextual factors may have a role in determining their desirability. Moreover, in the fourth chapter I analyse how non-cognitive values have influenced research in Evolutionary Psychology. Specifically, I analyse the influence of feminist values on the accounts of human mating suggested by Evolutionary Psychologists.

The second method I have adopted is the experimental method. In the fifth chapter, I will present and discuss the results of a survey aimed at exploring the role of non-cognitive values in cases of inductive risk. The experimental method has been only rarely used in the Science & Values debate, but it has recently started attracting more attention as a valuable method to clarify what the roles that values play and should play in science are.

I will present the two methods in the following sections. I will first present the fundamental assumptions of Evolutionary Psychology and I will then present the importance of experimental methods for the Science & Values debate.

2.1 Case Study: Evolutionary Psychology

John Tooby and Leda Cosmides have provided a list of the theoretical tenets of Evolutionary Psychology (Tooby & Cosmides 2005, 16-18). Here are the tenets as summarized in Downes 2018:

1. The brain is a computer designed by natural selection to extract information from the environment.
2. Individual human behaviour is generated by this evolved computer in response to information it extracts from the environment. Understanding behaviour requires articulating the cognitive programs that generate the behaviour.
3. The cognitive programs of the human brain are adaptations. They exist because they produced behaviour in our ancestors that enabled them to survive and reproduce.
4. The cognitive programs of the human brain may not be adaptive now; they were adaptive in ancestral environments.
5. Natural selection ensures that the brain is composed of many different special purpose programs and not a domain general architecture.
6. Describing the evolved computational architecture of our brains allows a

systematic understanding of cultural and social phenomena.

To begin, two main points should be noted. First, according to Evolutionary Psychologists, the great majority of the psychological mechanisms human beings currently possess are traits that have been selected over the course of human evolution (tenets 3 and 4), specifically in the environment of evolutionary adaptedness (EEA). EEA is “not a place or a habitat, or even a time period... [but] a statistical composite of the adaptation-relevant properties of the ancestral environments encountered by members of ancestral populations, weighted by their frequency and fitness-consequences” (Tooby & Cosmides 1990, 386–7). Despite claiming that EEA is not a place or era, Evolutionary Psychologists usually refer to the era of Pleistocene (a period of time approximately from 1.8 million to 10,000 years ago) as the period in which human beings have evolved the psychological traits we can observe today.

Second, Evolutionary Psychologists think human psychological mechanisms solve specific adaptive problems, i.e., specific problems of survival or reproduction (tenet 5). Each psychological trait evolved by human beings has a specific function, i.e., it is dedicated to solve a certain type of adaptive problem. This means that many of the psychological traits of human beings are explained by Evolutionary Psychologists as adaptations:

“characteristic *c* is an adaptation for doing task *t* in a population if and only if members of the population now have *c* because, ancestrally, there was selection for having *c* and *c* conferred a fitness advantage because it performed task *t*.” (Sober 2000, 85)

where by “fitness advantage” is meant success in spreading one's own genes. However, traits can be explained by appealing to the various products of natural selection. For instance, traits can be explained as exaptations, which are traits that were selected to solve an adaptive problem but have been later co-opted to serve another role (Lloyd 2015; see also Gould & Vbra 1982). Moreover, traits can be explained as by-products, namely “characteristics that do not solve adaptive problems and do not have functional design; they are “carried along with characteristics that do have functional design because they happen to be coupled with those adaptations” (Buss 1999/2008, 39).

Evolutionary Psychologists are interested in understanding the evolutionary history of psychological traits. This means that psychological mechanisms may be explained as adaptations, exaptations or by-products. However, the method used by Evolutionary Psychologists to study psychological traits is based on the assumption that traits are adaptations. That is, they apply what in the literature is called *methodological adaptationism* (which played a role in the female orgasm research presented in the first pages of this introduction), which can be explained as follows:

“the best way for scientists to approach biological systems is to look for features of adaptation and good design. Adaptation is a good “organizing concept” for evolutionary research.” (Godfrey-Smith 2001, 337)

Looking for adaptations, then, is the best heuristic to extend our body of knowledge on the evolutionary origin of human psychology. On the basis of this methodological assumption, Evolutionary Psychologists use two main strategies to formulate hypotheses concerning the evolutionary history of a trait. First, they consider a known trait of human psychology and suggest an evolutionary hypothesis explaining why that specific trait was selected, i.e., what was the function solved by that trait in EEA (reverse engineering). Second, they hypothesize a problem faced by our ancestors in EEA and argue that a certain trait was evolved to solve that specific problem (adaptive thinking). These strategies, or heuristics, will be analysed in detail in the second chapter on the value of fruitfulness.

Evolutionary Psychologists argue they do not endorse other kinds of adaptationism. For instance, they would not endorse *empirical adaptationism*, namely the idea that natural selection has caused most – if not all – of the biological and psychological traits of human beings. Empirical adaptationism is then a claim about the actual composition of the world: adaptations are the primary, omnipresent component of the biological world. In this view, while natural selection has a tremendous causal importance in evolution, other evolutionary factors have a limited role (Godfrey-Smith 2001, 336). Natural selection's activity would then meet few constraints, and factors as developmental and environmental constraints would not seriously inhibit the production of optimal traits.

Moreover, Evolutionary Psychology would not be committed to *explanatory adaptationism* either, according to which explaining the biological world by appealing to natural selection is “the core intellectual mission of evolutionary theory” (Godfrey-Smith 2001, 336). In other words, explanatory adaptationism ascribes different levels of importance to different kinds of evolutionary explanations. In particular, it attributes primary importance to explanations in terms of adaptation. Such a preference for these explanations is a matter of aesthetic: explanations in terms of adaptations have a special, privileged status because, in this view, they are more interesting and fascinating than non-adaptationist ones.

Evolutionary Psychology has been the target of many criticisms. For instance, its approach has been criticized as strongly committed to empirical adaptationism. Evolutionary Psychologists, in fact, make claims that call into question the idea that Evolutionary Psychology is only committed to the methodological version of adaptationism. For instance, Buss writes:

“all sexual psychological mechanisms owe their existence to evolution by selection” ... “evolution by selection is the only causal process powerful enough to produce complex organic mechanisms, including sexual mechanisms.” (Buss 1998, 23)

Robert Richardson (2007) has argued that Evolutionary Psychologists claim that they acknowledge the limits of natural selection and know that adaptations are not optimally designed (see for instance Buss et al. 1998, 539). However, their actual approach, Richardson argues, is not consistent with these modest claims and it would endorse a strongly adaptationist view involving both empirical and methodological adaptationism (Richardson 2007, 59). The problem with such a strong adaptationist view is that it may make it difficult to recognize that some traits are not adaptations. This issue will be discussed in the second chapter.

Some philosophers of science have also questioned the ability of Evolutionary Psychology to formulate good evolutionary hypotheses and gather evidence in support of them. As mentioned, one of the strategies used to formulate hypotheses is adaptive thinking, which involves hypothesizing a problem faced by our ancestors in EEA and arguing that a certain trait evolved to solve that specific problem. David Buller has argued that since we do not know the specific adaptive problems faced by our ancestors in EEA (Buller 2005, 94), Evolutionary Psychologists find themselves in the uncomfortable position of having little or no basis to formulate hypotheses about the specific traits evolved in human beings.

Evolutionary Psychologists use various bodies of evidence as a basis for the formulation of evolutionary hypotheses, such as the knowledge we have on current hunter-gatherer communities and on species related to ours, namely non-human primates. However, Buller argues these are unreliable sources of evidence. The current hunter-gatherer populations are taken as a source of evidence by Evolutionary Psychologists because our ancestors lived in hunter-gatherer communities. Hence, studying the adaptive problems faced by current hunter-gatherer populations could reveal the adaptive problems faced by our ancestors. However, Buller argues that “it is naive to think that the social lives of extant hunter-gatherer populations have not changed significantly in the last 10,000 years” (Buller 2005, 95). The problems faced by current hunter-gatherer communities are not necessarily the same problems faced by our ancestors in EEA.

There are also problems in using primates as source of evidence to formulate hypotheses. Buller argues that the most recent ancestor common to human beings and non-human primates (namely chimpanzees) lived 5 to 7 millions of years ago, and from that point the two species may have diverged significantly (Buller 2005, 96). The adaptive problems faced by non-human primates may then be not representative of the problems faced by our ancestors in EEA.

Critiques of Evolutionary Psychology have touched other points, such as the testability of the theories it proposes (Lewontin 1998; Richardson 2007) and the fact that research in Evolutionary Psychology would be influenced by sexist and androcentric assumptions (e.g. Lloyd 2003 on the evolutionary explanation of rape), problem that, as already mentioned, will be explored in the fourth chapter of this dissertation. These

assumptions would undermine the ability of the program to provide an accurate understanding of the evolutionary history of human beings' psychology.

However, Evolutionary Psychology's commitment to adaptationism and its ability to accurately formulate and test hypotheses on the evolutionary origin of psychological traits have also been defended by scientists as well as philosophers of science (e.g. Machery 2011; Lewis et al. 2017). For instance, Timothy Ketelaar and Bruce Ellis defend Evolutionary Psychology by recurring to the Lakatosian model of science (Ketelaar & Ellis 2000). They do so by referring to Imre Lakatos' idea of progressive research program, namely a program able to provide hypotheses and predictions that are corroborated by the evidence collected and able to digest anomalies (Lakatos 1970). By taking the theories on family violence, jealousy, and parental investment formulated by Evolutionary Psychologists as examples, Ketelaar and Ellis argue Evolutionary Psychology would be a progressive research program, providing novel hypotheses and predictions that have been successfully corroborated by evidence (Ketelaar & Ellis 2000, 13).

Evolutionary Psychology constitutes an interesting case study for the Science & Values debate for various reasons. One is that there is disagreement over the scientific status of the field. Some argue Evolutionary Psychology is bad science, facing fundamental evidential problems and using an unreliable methodology (namely methodological adaptationism), while others have defended the field as a promising approach to the study of human psychology (e.g. Machery 2011). Analysing Evolutionary Psychology by using cognitive values can shed new light on this debate. In the second and third chapters of this dissertation, I will employ the values of fruitfulness, explanatory power, and external consistency to explore some of the issues introduced in this section.

Moreover, Evolutionary Psychology constitutes an interesting case to study the influence of non-cognitive values on science. As mentioned, in the fourth chapter I will analyse the roles that feminist values played in the critique of the research program. However, Evolutionary Psychology is congenial to introduce a further issue concerning the influence of non-cognitive values on science, namely non-cognitive values' impact on the public understanding of science.

Evolutionary Psychologists have formulated evolutionary explanations for a wide array of phenomena, such as fears and phobias (e.g. Nesse 1990), homicidal fantasies (e.g. Duntley & Buss 1998), causes of divorce (e.g. Betzig 1989), poisonous animal aversion (Öhman et al. 2001), and disgust for certain food (Tybur et al. 2009). The research conducted on human mating has been particularly prolific.

Evolutionary Psychologists have conducted studies on various traits related to the romantic and sexual life of human beings, such as studies on jealousy (e.g. Buss et al. 1990), sexual fantasies (e.g. Ellis & Symons 1990), and mate preferences (e.g. Singh 1993; Gangestad & Thornhill 1997). Many of these studies focus on understanding how women and men differ in their sexual behaviours. They show that women and men would differ in

the kinds of jealousy they experience (e.g. Buss et al. 1992), mates preferences (Buss et al. 1990), and the kinds of relationships desired (e.g. Buss & Schmitt 1993).

The studies on human sexual behaviours is the line of research that has attracted the most attention among the public, being the most discussed in mass media (Cassidy 2005). Cassidy 2005 has analysed the popularisation of the program in the UK in the 1990s and revealed that, in general, Evolutionary Psychologists' research had a considerable press coverage, with both Evolutionary Psychologists engaging in the public debate and non-experts writing about this research. Other fields did not receive the same attention.

The popularity of Evolutionary Psychology's studies on human sexual behaviours is not very surprising. These studies provide information useful for answering some questions that may be perceived as very important or fascinating by human beings, such as do men's and women's sexual lives differ in important respects? Do we have any chances to find a long-term committed partner? Are we evolutionary well-equipped to attract the persons we find interesting? Moreover, one of the reasons of this popularity may be that Evolutionary Psychology's findings easily fit with widespread stereotypes, such as the findings over the preference of men for younger women (Greenlees & McGrew 1994) and the preference of women for men with status and resources (Buss et al. 1990). In the fifth chapter, I present a study on the effects that non-cognitive values can have on public reception of scientific research. In particular, I will examine the relation between non-cognitive values, personal features, and people's reactions to cases of scientific uncertainty. The next section will clarify the methods used to conduct this study.

2.2 Experimental Philosophy

The second strategy I have used in this dissertation is the experimental method. Experimental philosophy aims at exploring philosophical issues by means of experimental methods (e.g. surveys and linguistic studies) usually employed in psychology and social sciences. The idea is that empirical studies can be used to address philosophical questions (e.g. Knobe 2003; Machery et al. 2015). These methods, experimental philosophers argue, can provide philosophers with information that the traditional conceptual analysis cannot offer. For instance, they can provide information relevant to clarify the empirical premises of philosophical arguments and to reveal the influence of biases on philosophers' intuitions about certain thought experiments (O'Neill & Machery 2004).

To date, the Science & Values debate has been based on case studies and the use of experimental methods is still limited. For instance, Colombo and colleagues have focused on the moral offensiveness of scientific hypotheses (Colombo et al. 2016). Their results show that judgments about the moral value of scientific hypotheses can influence people's explanatory judgments, namely whether hypotheses are seen as supported by evidence or not. Moreover, Steel and colleagues conducted a study on scientists' attitudes towards non-

cognitive values, aimed at understanding which roles, in scientists' opinion, non-cognitive values have and should have in science (Steel et al. 2017). Scientists were asked questions concerning various issues, such as the nature of scientific enterprise (e.g. whether value-free science is possible), and the responsibility of scientists (e.g. whether researchers should engage in the promotion of social aims, such as public health) (Steel et al. 2017, 24). Their results reveal gender differences in scientists' views about roles of values in science and that scientists do not "appear to favor the value-free ideal" (Steel et al. 2017, 25).

The experimental method can provide information relevant to the Science & Values debate that cannot be gathered or easily gathered through the analysis of episodes of the history of science, such as information on the ideas scientists have on the roles of values in science and what they think about their moral and social responsibilities (Steel et al. 2017, 26). The case study method has also been criticized as limited in its reliability and the conclusions it allows. The method would be unreliable since philosophers may choose cases from the history of science so as to fit their predetermined philosophical positions. The analysis would then be significantly biased. Moreover, the generalizability of the conclusions drawn from a case study may be limited. That is, what a case study shows, say, on the influence that values have in a specific episode of the history of science, is not necessarily representative of the influence that values - in general - have on science (Steel et al. 2017, 23; see also Pitt 2001).

However, also experimental methods face some problems. For instance, surveys has been criticized as of limited importance for the philosophical discussion because providing collections of "quick, unreflective judgments, which are not an appropriate source of evidence about concepts" (O'Neill & Machery 2004, ix).

This dissertation will show that the Science & Values debate can benefit from the use of both methods, which are complementary in the type of information they provide. While I will use the case study method to understand how cognitive values can be applied in theory appraisal and the epistemic benefits that non-cognitive values can provide in science by appealing to examples from the history of science and current scientific practice, I will use the experimental method to understand how non-cognitive values can affect human psychology.

In the fifth chapter, I will present a study aimed at clarifying the roles of non-cognitive values in scenarios involving inductive risk. The philosophical literature has provided several case studies on the importance of non-cognitive values in cases of inductive risk (e.g. Douglas 2000; Staley 2017), but little attention has been paid to how people actually reason about uncertainty in science and possible errors. However, understanding the psychology of inductive risk has philosophical relevance. In particular, the results of the study will be used to explore questions over the value-free ideal and the role that citizens' non-cognitive values can legitimately play in science, in the light of the recent calls made by institutions to encourage the engagement of citizens into science (Figueiredo Nascimento et

al. 2016; Funk et al. 2015) and the alignment of scientific research with the needs and values of citizens.

3. Summary and final remarks

In this introduction, I have specified the questions I address in this dissertation, their theoretical background, and the methods I use to explore these questions.

The dissertation can be divided in two main parts. While the second and third chapters focus on questions concerning the role and importance of epistemic and cognitive values in science, the fourth and fifth aim at clarifying the roles of non-cognitive values in science. However, the chapters share an underlying aim, namely to show that the value-free ideal of science is not a desirable ideal. Epistemic and cognitive values constitute useful tools to assess and compare scientific theories and non-cognitive values may significantly promote the attainment of knowledge. Moreover, they aim at showing that the value-free ideal is not a realistic portrayal of science. As the studies presented in the fourth and fifth chapter reveal, non-cognitive values play a role in the history of science and constitute an important component of how people reason about uncertainty in science.

In this dissertation I have used Evolutionary Psychology to understand the roles and importance of values in science. This means that the claims I make about epistemic, cognitive, and non-cognitive values might have limited scope, i.e., they might be non-generalizable to other fields or research programs. However, this should not be seen as a problem, but a feature of my argument. After all, as I will argue in Chapter 3, any analysis of the roles and importance of values in science should be *local*. This is because the roles and importance of values can significantly vary across contexts and an accurate understanding of the roles of values should then take into account this variance.

Chapter 2

What we (should) talk about when we talk about fruitfulness¹¹

1. Introduction

There is not a general consensus about the definitions and roles of several cognitive values¹² included in the lists suggested by philosophers (e.g. Kuhn 1977; McMullin 1983; Douglas 2000; see also section 1.1 of the introduction). While many papers have shed light on some of these values, such as simplicity (e.g. Forster & Sober 1994; Baker 2003), little attention has been paid to other values, such as fruitfulness. Few studies have addressed the explication of fruitfulness as a value relevant to assess and compare rival scientific theories or research programs. Much of the literature provides definitions that are too vague for providing strategies to assess the fruitfulness of theories and programs. A good illustration of this problem is Kuhn's account of fruitfulness. Despite ascribing a prominent role to fruitfulness in theory choice, he spent very few words on it. When defining it, he just states that fruitful theories "disclose new phenomena or previously unnoted relationships among those already known" (Kuhn 1977, 322). Moreover, there is considerable disagreement over the importance of fruitfulness in theory assessment. While some philosophers argue that fruitfulness is a crucial value for the assessment of theories and programs (McMullin 1976; Kuhn 1977), others claim that it plays a secondary role in theory appraisal (Nolan 1999; Douglas 2009). For instance, Hugh Lacey (1999) argues that fruitfulness should not be

¹¹ This chapter is a revised version of Ivani (2019).

¹² If not specified otherwise, in this chapter I will use "cognitive values" to refer to both epistemic and cognitive values.

included in the list of values that are relevant for assessing theories and programs. According to Lacey, fruitfulness “pertains more to such stances as the provisional entrainment of theories rather than to their acceptance” (Lacey 1999, 60). Fruitfulness has also been criticized as having a noxious influence on science. Helen Longino (1995) criticizes the Kuhnian notion of fruitfulness because it focuses on the development of the content of theories in response to puzzles that have to be solved. Instead, the development of a program should be assessed as valuable, Longino argues, if it “can be used to improve living conditions in a way that reduces inequalities of power” (Longino 1995, 394).¹³

Overall, despite being frequently mentioned by philosophers and scientists, how to understand fruitfulness remains obscure. A satisfactory explication should help articulate concrete strategies for assessing and comparing the fruitfulness of research programs and should then answer at least two questions: What is fruitfulness? What are reliable strategies for assessing and comparing the fruitfulness of rival research programs in science? The aims of this chapter are to provide an explication of fruitfulness and a strategy to assess it.

Despite being vague, the literature on fruitfulness provides useful insights for developing an in-depth study. Many of the accounts suggested so far share the core idea that fruitfulness would concern the desirable development involving changes, extensions, and improvements of research programs’ content. In this view, fruitful programs can produce epistemic goods. I argue that in order to provide a satisfactory explication of fruitfulness, this core idea should be maintained and developed further. In particular, I suggest assessing fruitfulness by taking into account the set of tools that makes the development possible. This set is what scientists can employ for developing the content of the program. These are the tools for formulating and testing hypotheses, such as experiments and methods to analyse data. I argue that focusing on two of these tools, namely the research questions and discovery heuristics, provides us with a concrete, simple strategy for evaluating and comparing the fruitfulness of research programs. Considering research questions and discovery heuristics constitutes a novel approach to fruitfulness. To date, the literature has failed to recognize these factors as key elements for the assessment of fruitfulness.

This chapter is structured as follows. In the second section, I consider the literature on fruitfulness and I illustrate the aforementioned core idea as presented by Thomas Kuhn, Ernan McMullin, Daniel Nolan, and Samuel Schindler. Furthermore, I outline my approach to assess fruitfulness, which is focused on research questions and discovery heuristics. In the third section, I introduce a case study, namely the adaptationist program in Evolutionary Psychology. This case study shows how considering questions and heuristics can clarify and simplify the assessment of fruitfulness of research programs. A remark on the terminology I use: talking about ‘research programs’ recalls Imre Lakatos’ philosophy. However, I am not

¹³ Longino also suggests replacing fruitfulness with two alternative values, namely applicability to current human needs and diffusion of power (Longino 1995, 389; see also section 1.1 of Chapter 1).

committed to the specific aspects of the Lakatosian notion of research program, such as hard-core and protective belt (see for instance Lakatos 1970). I construe a research program as a set of theories that share some fundamental assumptions and methodologies. In the fourth section, I draw conclusions.

2. What is fruitfulness?

Scientists frequently refer to fruitfulness as an important feature of research programs. Max Planck talks about the fruitfulness of explaining entropy by referring to the probability calculus in physics (Planck 1915, 35). Henri Poincaré praises the fruitfulness of mathematical reasoning (Poincaré 1905, xix). The Evolutionary Psychologist David Buss stresses how research on sex differences has become more fruitful by focusing on (internal) psychological mechanisms rather than on (external) manifest behaviours (Buss 1998, 22). In particular, Buss claims that this shift was fruitful because it led to the formulation of several hypotheses and the discovery of many sexual differences. Stephen Gould, instead, criticizes evolutionary psychology as fruitless. He claims: “evolutionary psychology could, in my view, become a fruitful science by replacing its current penchant for narrow, and often barren, speculation with respect for the pluralistic range of available alternatives” (Gould 1997). There are many similar examples that could be cited.

However, the pervasiveness of fruitfulness in scientific papers is not mirrored in the philosophical literature. Relatively little attention has been paid to explicating fruitfulness and its role in theory appraisal. One of the consequences of the lack of a clear explication is that fruitfulness can be virtually ascribed to many different programs, even rival ones. The fact that we can ascribe fruitfulness to rival programs does not constitute a problem. Indeed, two rival programs can be both fruitful, just as they can be both simple or internally consistent. However, the problem with fruitfulness is that it can be easily ascribed to many programs because its definition is loose and no clear strategy for assessing it is provided.¹⁴ As the above examples of Buss and Gould show, scientists may radically disagree whether a program is fruitful or not. This may depend on the divergent ideas that scientists entertain about the merits of a research program. However, in the next sections I argue that such disputes may be fostered by a vague definition of fruitfulness and that it is possible to shed light on them by appealing to a clear account of fruitfulness.

In the following pages, I explicate a notion of fruitfulness understood as a capacity to develop of programs. However, before doing that, in the next subsection I discuss the accounts suggested by Thomas Kuhn, Ernan McMullin, Daniel Nolan, and Samuel Schindler.

¹⁴ This problem affects many values. As mentioned in the introduction, several papers discuss how to best define simplicity (e.g. Forster & Sober 1994; Baker 2003). It is beyond the scope of this chapter to address the problems arising from the vague definitions of other values.

While Kuhn and McMullin provide two of the earliest attempts to explicate fruitfulness, the accounts of Nolan and Schindler are two of the rare recent analyses of fruitfulness.

2.1 Fruitfulness as desirable development of programs

Thomas Kuhn includes fruitfulness in his pioneering list of values (1977). Although he claims that fruitfulness is “of special importance to actual scientific decisions” (Kuhn 1977, 322), he spent very few words on this value. Nonetheless, from the little he says, it is clear that he construes fruitful programs as producing new research findings and disclosing “new phenomena or previously unnoted relationships among those already known” (ibid.). In particular, fruitful programs produce new findings and discoveries that help to reach solutions to research puzzles. For instance, Kuhn claims that the adoption of mechanico-corpuscular explanations in different programs made them fruitful since it led to the solutions of “problems that had defied generally accepted solution and suggesting others to replace them” (Kuhn 1962, 104).

Fruitfulness plays a crucial role in the Kuhnian picture of theory choice, influencing scientists’ decision leading to transitions from old paradigms to new paradigms.¹⁵ Paradigms are, in Kuhn’s view, puzzle-solving exemplars, determining the rules for solving puzzles and what counts as an acceptable solution (Kuhn 1962, 38). When a paradigm and its rules persistently fail to solve puzzles, other paradigms and new rules emerge and scientists become interested in these other programs. Hence, the power of solving puzzles, i.e. the fruitfulness of programs, is one of the main reasons that convince scientists to work with a paradigm.

Kuhn’s ideas have influenced other philosophers’ understanding of fruitfulness. For instance, Ernan McMullin elaborates the idea of fruitfulness as the power of programs to develop and solve puzzles and anomalies. Indeed, in his view “a fertile theory is one with the capacity to grow and change” (McMullin 1979, 60) by finding solutions to puzzles and anomalies.¹⁶ Such a capacity is a diachronic virtue of programs, i.e., it manifests itself over the course of time (McMullin 2009, 505). To account for this time-dependent aspect, McMullin proposes two kinds of fruitfulness, namely proven fertility and untested fertility. While proven fertility is past oriented and it concerns the past success of a program in growing and changing, untested fertility is future oriented and refers to the potential ability to develop in the future (McMullin 1976, 400). McMullin does not ascribe equal importance to the two kinds of fruitfulness. Indeed, he argues that while proven fertility is a truth-

¹⁵ While Kuhn mentions theories and paradigms in his analysis of fruitfulness, I talk about programs. Although these are not synonyms, their differences do not constitute a problem for my analysis of fruitfulness.

¹⁶ McMullin employs the term “fertility” to refer to the virtue discussed in this chapter. In the literature, philosophers also use “fecundity” (Quine 1955) and “boldness” (Colyvan 1999).

indicative feature of programs, untested fertility is not and, because of this, focuses his analysis on proven fertility.¹⁷

McMullin explicates fruitfulness as follows:

“The [fruitful] theory proves able to make novel predictions that were not part of the set of original explananda. More important, the theory proves to have the imaginative resources, functioning here rather as a metaphor might in literature, to enable anomalies to be overcome and new and powerful extensions to be made. Here it is the long term proven ability of the theory or research program to generate fruitful additions and modifications that has to be taken into account.” (McMullin 1983, 16)

In this view, fruitful programs have the adequate instruments to promote the flowering of new ideas and extensions. They stimulate the creative imagination of scientists to modify the content of programs in desirable ways (McMullin 1979, 60). As Robert Segall explains, according to McMullin fruitful programs possess metaphorical suggestiveness: just like good metaphors, fruitful programs suggest creative ideas to improve our understanding or knowledge of the world (Segall 2009, 58). However, being fruitful does not mean encouraging mere change. A type of change, McMullin argues, is of special importance. He suggests focusing on the modifications of a program occurring when it has to deal with anomalies, which he calls non-*ad hoc* accommodations. These modifications are valuable since they involve adjustments that do not modify the fundamental ideas of the program (McMullin 1985, 264). By way of illustration, McMullin discusses the discovery of the electron spin in the Bohr model. Scientists were not able to explain some anomalies, i.e., the reaction of hydrogen to a strong magnetic field. Then, such anomalies were successfully accommodated by postulating electron spin. McMullin claims:

“Electron spin was not part of the original Bohr model. Its omission was not a formal idealization, strictly speaking, because there was no reason to suppose the electron would possess spin. ‘Has the electron spin?’ was a question that simply had not been asked, and that could not be answered within the original model. But it was a question that at some point could easily enough suggest a fruitful line to follow.” (McMullin 1985, 263–4)

For McMullin, this accommodation was inspired by the theory itself and it harmoniously fitted with the program in its original formulation. The intrinsic creative suggestiveness of the program produced this adjustment. This ability to suggest modifications, McMullin argues, is a truth-indicative virtue of programs, i.e., it reveals that the content of the program is true or approximately true. That is, the ability of a program to survive to anomalies by

¹⁷ The focus on proven fertility was dictated by McMullin’s purpose of employing fruitfulness to defend scientific realism (McMullin 1976).

means of changes that extend and improve its content reveals that the program gets something right about the real structure of the world.¹⁸

Daniel Nolan has criticized McMullin's account of fruitfulness. Specifically, he claims that proven fertility can be reduced to other values and untested fertility's importance is parasitic on the importance of other values. Indeed, a program has proven fertility if it is able to confirm itself through changes and this, Nolan argues, is a facet of the value of novel confirmation (Nolan 1999, 271). On the other hand, he construes untested fertility as the chance that a future version of a program will perform better in fulfilling some virtues than its current version, such as strength and predictive power (Nolan 1999, 271). Still, despite being a dismissive, reductionist account, Nolan's portrayal shares the core idea of fruitfulness as the power of programs to develop with Kuhn and McMullin's accounts. Kuhn and McMullin construe fruitfulness as an ability of programs to develop by finding solutions to puzzles and anomalies. Nolan's account, instead, focuses on the desirable development involved in novel confirmation and fulfilling certain virtues in the future. More recently, McMullin's analysis of fruitfulness has been criticized by Samuel Schindler (2017). In particular, Schindler argues that the discovery of the electron spin is not a genuine case of non-*ad hoc* accommodation (Schindler 2017, 168). Although Schindler criticizes this historical case study, he claims that McMullin's notion of fruitfulness should be developed further. In particular, the relation between fruitfulness and non-*ad hocness* should be studied since, Schindler argues, non-*ad hocness* clearly is an important virtue of programs (Schindler 2017, 170). Nolan and Schindler provide the most recent accounts of fruitfulness. Although many recent studies have analysed the role of values in scientific practice (e.g. Douglas 2009; Steel 2010), they have largely ignored the literature on fruitfulness. In the next section, I develop the core idea of these accounts to provide an explication of fruitfulness and a strategy to assess programs.

2.2 Developing the notion of fruitfulness

Let's define fruitfulness as *an ability of research programs to develop*. A program develops when it enhances our knowledge or understanding of the world. In particular, fruitful programs develop by producing epistemic goods, such as novel hypotheses. Now, how can we assess this capacity to be productive? What are the relevant aspects for evaluating the productivity of programs? To answer these questions, I suggest focusing on the set of tools of research programs. This is what the researchers can use to extend and improve the content of a research program. Not any extension of the content counts as an improvement. For

¹⁸ Recently, Schindler (2017) has pointed out that the non-*ad hoc* modifications analysed by McMullin are akin to Imre Lakatos' positive heuristic (Lakatos 1970). They both involve adjustments that do not seriously alter what Lakatos calls the "hard core" of the program, but they are able to extend its content. Schindler calls these modifications "de-idealizations of the theory" (Schindler 2017, 156).

instance, making the content broader by formulating many implausible hypotheses does not make the program developing. The *quality* of the hypotheses generated should also be considered. Specifically, I suggest analysing two tools that scientists employ to extend programs' content, namely *research questions* and *discovery heuristics*. These components are used to organize and plan research and to formulate novel hypotheses. They guide and constrain – together with other components – the development of research programs. Both the kinds of questions and heuristics and the ways in which they are employed are relevant factors to analyse for assessing the fruitfulness of programs.

Consider research questions first. Several studies have explored the importance of questions in the process of extending our knowledge (e.g. van Fraassen 1980; Mason 2002). For instance, some philosophers suggest understanding science as a question-answering process (Hintikka 1981). The questions we ask show what we find worthy of being investigated or what we think as a useful starting point for extending our knowledge. Examples of research questions are:

- (a) What is the evolutionary function of jealousy for human beings?
- (b) Is there a link between flu vaccines and miscarriage?

Now, why are research questions particularly relevant for assessing the fruitfulness of research programs?

Research questions constitute the starting point of the enquiry and, as Sven Lundstedt points out, “initial scientific questions, like first impressions, carry a great deal of weight in shaping the direction of a system of thought” (Lundstedt 1968, 229). The research questions posed and the answers that they allow as appropriate significantly affect the development of research programs: questions can either promote or compromise the fruitfulness of programs.

Research questions have an important role in determining the methods used to conduct the research, which sizes of samples are legitimate, and what counts as relevant data (Mason 2002, 19).¹⁹ The choice of research questions depends on several factors, such as the background assumptions endorsed by the scientists or the scientific community. These assumptions can be of different sorts, e.g. moral, political, and methodological, and their importance can vary across contexts. For instance, question (b) may be the result of some moral assumptions, such as the idea that protecting human health is a moral duty and our research should then be focused on preventing damages caused by medicines. Such a moral concern may play little or no role in determining the question over the function of jealousy.

Analysing research questions can provide us with relevant information about the research conducted. Specifically, research questions have at least three dimensions that can

¹⁹ See Bryman (2007) for a critique of the idea of research questions as fundamental factors for research design.

be explored. First, examining a research question makes it possible to understand the *focus* of the research, i.e., the portion of the world that scientists want to investigate. Question (a) makes clear that human beings and their psychological traits are the phenomena on which the research primarily focuses. Second, research questions show the *aims* of the study. For instance, question (b) makes clear that the goal of such an investigation is gaining knowledge on the possible harms caused by some vaccines. Finally, by analysing research questions we can understand the *orientation* of the enquiry. Namely, research questions provide the direction of the investigation, i.e., the track along which the research can develop. Questions guide the choice of methodologies for formulating and testing hypotheses: the appropriateness of a method depends on the chances that such a method has of answering the question. Indeed, when it comes to answering a research question, some methods are appropriate, some others are not. To illustrate, let's consider question (a). The aim of the study is investigating the function of jealousy and surveys and cross-cultural studies constitute – among others – appropriate methods for finding an answer. On the other hand, it is unlikely that using a cloud chamber helps to find an answer to this question.

In addition to constraining the choice of methods, research questions determine what a legitimate answer is. Here is where the relation between research questions and fruitfulness becomes clear. Not any sort of answer is good as a reply to a question. For instance, the answer “Jealousy is one of the causes of many stalking cases” to the question “What is the function of jealousy in evolution?” does not constitute a proper answer. That answer fails to address the specific aim of the question, namely clarifying the evolutionary history of jealousy. However, there is more to be discussed about the pertinence of answers. In a recent paper, Elisabeth Lloyd clarifies this point. She claims that “asking different questions makes contrasting classes of answers legitimate” and “each question carries with it an appropriate class of possible answers unique to it, and distinct from other contrasting classes of answers” (Lloyd 2015, 346). This means that research questions make some specific answers appropriate and some others irrelevant. As a consequence, while some answers are considered and analysed, others are ignored. This may negatively affect the development of research programs. Indeed, as Lloyd argues, a problematic use of research questions “can lead us to miss what's really going on, therefore to scientific failure” (Lloyd 2015, 346). By excluding some answers, questions can constrain research and have an impact on the effective development of research programs. I argue that some questions may perform better than others when it comes to making a program fruitful. In the next section, a case study illustrates how the choice of research questions and the specific way in which they are used may either undermine or stimulate the development of programs.

Besides research questions, a good assessment of the fruitfulness of research programs should consider a further component of their set of tools, namely the discovery heuristics. These are the strategies used to formulate novel hypotheses. Heuristics guide, constrain, and help scientists' imagination in the exploration of phenomena. They are not

perfectly reliable strategies affording scientists good hypotheses in any context. However, they are taken as useful techniques to approach the world (Wimsatt 2007, 76). Assessing the discovery heuristics of a program, then, means evaluating its ability to produce epistemic goods.

What are the discovery heuristics that can make a program fruitful? Fruitful programs use *reliable* heuristics. That is, the pattern suggested by the heuristics for formulating hypotheses is rigorous, i.e., it does not involve defective reasoning. Reliable heuristics allow scientists to formulate *well-designed hypotheses*. This does not mean that reliable heuristics only produce hypotheses that turn out to be true. Rather, well-designed hypotheses are testable hypotheses that have good chances to be true. Moreover, it should be noted that heuristics producing many novel hypotheses do not necessarily make a program fruitful. Proliferation of hypotheses is not a mark of fruitfulness. Rather, it may indicate the sloppiness of programs. The strategy for evaluating fruitfulness suggested in this chapter makes it possible to distinguish between fruitful programs and “creative-but-stagnating” programs. It permits to draw a distinction between virtuous programs, whose content grows and improves over time, and vicious programs, whose content may grow by producing many hypotheses, but does not improve, i.e., programs that do not extend our understanding of the world. As mentioned, proliferation of hypotheses does not necessarily constitute desirable development. This point will be discussed in detail in the next section.

To summarize, in this section I have presented the main accounts of fruitfulness and I have introduced my explication of fruitfulness and a strategy to assess programs. In the next pages, I provide an example unveiling the efficacy of this strategy.

3. Fruitfulness at work: A case study

The adaptationist program is the leading research program in evolutionary psychology. It is led by the Santa Barbara School, which involves David Buss, John Tooby, Leda Cosmides, and many others. As I have mentioned in the introduction of this dissertation, in order to distinguish this program from other programs, I refer to this approach by using David Buller’s notation (2005) “Evolutionary Psychology” with capital letters.²⁰ Evolutionary Psychology is committed to *methodological* adaptationism (see section 2.1 of Chapter 1), which is the idea that looking for adaptations is the most efficient strategy to extend our body of knowledge. Methodological adaptationism does not necessarily involve the assumption that adaptation is the most widespread product of evolution. In other words, adaptationist

²⁰ Machery and Clark Barrett (2006) argue that any distinction made between the projects conducted in evolutionary psychology is problematic because they often share methodologies and topics. It is beyond the scope of this chapter to discuss this issue. However, the research discussed in this chapter is the one committed to methodological adaptationism.

programs are not necessarily committed to empirical adaptationism, i.e., the idea that natural selection has caused most – if not all – of the psychological traits. In this view, adaptations are the primary, omnipresent component of the biological world (Godfrey-Smith 2001, 336; see also section 2.1 of Chapter 1). Methodological adaptationism, instead, does not tell anything about the actual composition of the biological world. Its focus is the best way to explore it. Starting out the research by focusing on adaptations is, according to methodological adaptationists, the most effective and reliable approach to understand the evolutionary origin of traits, regardless of the actual amount of adaptations.²¹ As Tooby and Cosmides explain, methodological adaptationism is essential to Evolutionary Psychology:

“if one is interested in uncovering intelligible organization in our species-typical psychological architecture, discovering and describing its adaptations is the place to begin.” (Tooby & Cosmides 1992, 55)

Adaptationist psychologists have proposed plenty of hypotheses to explain a wide range of traits. They often stress the productivity of their approach. The praise of adaptationism’s ability to account for previously ignored phenomena and make discoveries is a *leitmotif* found in many papers (e.g. Lewis et al. 2017; Buss & Schmitt 2011). For instance, Stephen Pinker claims that “in the study of humans, there are major spheres of human experience – beauty, motherhood, kinship, morality, cooperation, sexuality, violence – in which evolutionary psychology provides the only coherent theory” (Pinker 2002, 135). Buss and Reeve mention this exceptional virtue of their approach:

“By any reasonable standard, evolutionary psychology has discovered an impressive array of empirically documented phenomena that were not discovered by prior mainstream nonevolutionary psychologists.” (Buss & Reeve 2003, 849)

To prove that, Buss and Reeve provide a long list of phenomena explained by Evolutionary Psychology, such as desire for sexual variety, stepchild abuse, and causes of divorce (Buss & Reeve 2003, 850).

Prima facie, Evolutionary Psychology seems to be an extremely fruitful research program. It is able to develop its content by suggesting many hypotheses that improve our understanding of the evolutionary origin of psychological traits. Evolutionary Psychology is portrayed as a constant blooming of new hypotheses and methodological adaptationism looks like an exceptionally fruitful strategy. Psychologists claim that their approach is so fruitful because it constitutes the most reliable and non-arbitrary way to classify and understand

²¹ Methodological adaptationism comes both in a strong and a weak form. While strong methodological adaptationism sees the adaptationist strategy as the only successful approach for understanding evolution, weak methodological adaptationism pictures it as just one of the successful strategies (see Lewens 2009, 169).

phenomena (Buss 1995, 6). Moreover, they argue that alternative approaches to the adaptationist program do not equal such a great triumph in extending our knowledge. In their view, this program has clearly shown its complete dominance in developing evolutionary psychology (Buss & Reeve 2003, 852).

Evolutionary Psychology has also been criticized because of its astonishing productivity. Many scholars challenge the overuse and soundness of adaptationist explanations (e.g. Richardson 2007). In their view, such a large amount of hypotheses would be a sign of the unreliability and sloppiness of methodological adaptationism.

So, is Evolutionary Psychology fruitful? How can we deal with this criticism? Let's assess Evolutionary Psychology's fruitfulness by analysing its research questions and discovery heuristics.

3.1 Research questions in Evolutionary Psychology

What are the research questions guiding Evolutionary Psychology? Buss provides us with an answer. The questions are:

“What is the nature of the psychological mechanisms that evolution by selection has fashioned? Why do these mechanisms exist in the form that they do – what adaptive problems did they arise to solve, or what are their functions?” (Buss 1995, 5)

The reference to natural selection and function makes it clear that these are adaptationist questions, which can be exemplified by the following question:

(AQ) What is the function of trait T?

Asking about the function of a trait means enquiring its evolutionary function. The question, then, revolves around adaptations. Following the aforementioned three dimensions of research questions, namely focus, aim, and orientation, it is possible to clarify some aspects of this question. The focus of the adaptationist question are biological traits. Its aim, then, is gaining knowledge on the evolutionary origin of these traits. Specifically, the question addresses the role of natural selection in the development of traits. Finally, the orientation of this question is adaptationist, i.e., the starting point is assuming that the trait studied is an adaptation.

Psychologists see asking this question as the most helpful procedure to advance our understanding of the role of evolution on psychological mechanisms. If the adaptationist question was not being used, many discoveries would not be possible (Cosmides & Tooby 1994, 43). However, in a recent paper Elisabeth Lloyd has criticized the use of adaptationist questions. She argues that such questions impoverish the quality of research in evolutionary biology and psychology. In her view, the adaptationist questions as used in Evolutionary

Psychology only admit adaptive answers, namely answers involving functions and adaptations (Table 2).

This means that this strategy rules out other possible answers, i.e., answers involving non-adaptive forces. Indeed, Lloyd claims that Evolutionary Psychologists see adaptive and non-adaptive answers as mutually exclusive. In particular, they see non-adaptive hypotheses as statistical nulls. They understand “the rejection of adaptive hypotheses as logically necessitated for consideration or acceptance of any nonadaptive hypothesis” (Lloyd 2015, 349). This is evident in the discussion of Evolutionary Psychology’s methodology offered by Buss and colleagues (Buss et al. 1998). They claim that when an adaptationist hypothesis is not supported by data, the following options are available to psychologists. First of all, researchers may want to check whether the tests have been done properly. Second, if the tests have been correctly done and the adaptationist hypothesis is just wrong, researchers can suggest alternative adaptationist hypotheses. Third, researchers may want to test non-adaptationist hypotheses. However, considering non-adaptationist hypotheses, Buss and colleagues argue, involves trying “to disconfirm all existing functional explanations” (Buss et al. 1998, 544). This – Lloyd argues – is misunderstanding how evolution works. Several evolutionary factors can be involved in the evolution of a trait, i.e., several answers can simultaneously account for a trait. The way Evolutionary Psychologists use the adaptationist question leads to ignore important evolutionary factors, such as by-products, that need to be considered to provide an exhaustive analysis of phenomena. As a consequence, the power of this program to provide adequate explanations is undermined.

Lloyd states that the fruitfulness of this inquiry improves with a different question: “What evolutionary factors may play a role in the form and structure of the trait T?” (Lloyd 2015, 346) (Table 3).

Non-adaptive and adaptive answers here are not seen as mutually exclusive. Namely, the question does not demand a rejection of adaptive answers in order to make room for non-adaptive ones. While the adaptationist question underestimates the actual role of different factors in evolution, the evolutionary factors question allows for different kinds of explanation and it is then able to better account for the complexity of evolution. The two questions differ in their orientations. The adaptationist question has an adaptationist orientation and sees assuming that a trait is an adaptation as the best strategy to study it. The evolutionary factors question, instead, does not make such an assumption and it involves a pluralistic methodology. This question makes it possible to offer a pool of stronger hypotheses that may enhance the fruitfulness of evolutionary psychology.

Question	Answers
What is the function of the trait T?	A: The function of this trait is F. A: The function of this trait is G. Etc.

Table 2. Adaptationist question and its answers (Lloyd 2015, 346).

Question	Answers
What evolutionary factors may play a role in the form and structure of the trait T?	A: This trait occurs in the population because it has the function F, i.e., the trait is an adaptation. A: This trait occurs in the population because it has the function G, i.e., the trait is an adaptation. A: This trait occurs widely in this population because it is genetically linked to a trait that is highly adaptive in this species (genetic linkage or correlation). A: This trait has its current form largely because of an ancestral pattern (phyletic inertia). A: This trait has its current form and distribution because of pleiotropy with a trait that was under natural selection (pleiotropy or by-product). A: This trait has its current form and distribution because it is a by-product or bonus of a trait that is strongly selected in the opposite sex in this species (by-product or bonus of an adaptation). A: This trait has its current form and distribution because of some combination of the above factors. Etc.

Table 3. Evolutionary factors question and its answers (Lloyd 2015, 346).

If Lloyd is right, the adaptationist question, as used in Evolutionary Psychology, may impede the desirable development of the program. In other words, the possibility of extending and improving the content is undermined by the use of a limiting question that ignores legitimate answers, namely answers that, because of the complexity of evolution, need to be considered. This means that by *monopolizing* the kind of answers pursued, the adaptationist question does not make Evolutionary Psychology a fruitful program. The use of this question may lead to what Philip Kitcher calls the “tyranny of the ignorant” (Kitcher 2003, 130). Kitcher argues that because of some moral, political background assumptions, some questions are often undervalued by scientists. The case study shows that important questions, such as the pluralistic one suggested by Lloyd, are ignored and this negatively affects the advancement of science. The tyranny of the adaptationist question in Evolutionary Psychology is not due to moral and political assumptions. Rather, it is attributable to the extreme confidence in the adaptationist approach as a reliable strategy to improve our understanding of the evolution of human psychology. That is, in this case the tyranny of the ignorant is due to methodological assumptions.

3.2 Heuristics in Evolutionary Psychology

The second factor I suggest for analysing the fruitfulness of research programs are the discovery heuristics, namely the strategies for developing novel hypotheses.

Evolutionary Psychologists apply two main heuristics to suggest hypotheses, i.e., reverse engineering and adaptive thinking (Griffiths 1996, 514; see also section 2.1 of Chapter 1). Reverse engineering “involves working backwards from what we know of human psychological architecture, to what its evolutionary functions might have been in ancestral environments” (Durrant & Haig 2001, 362). Scientists detect a human beings’ trait and imagine its selective advantage for our ancestors. For instance, if they have data about marital jealousy, psychologists formulate a hypothesis that explains such a trait as a solution to an adaptive problem faced by our ancestors in EEA. When using adaptive thinking, instead, researchers conceive “the adaptive problems ancestral humans were likely to have faced and generate plausible accounts for the psychological mechanisms which would have evolved to solve them” (Durrant & Haig 2001, 362). That is, scientists imagine a particular adaptive problem in EEA, such as protecting one’s own offspring, and they formulate a hypothesis concerning the existence of a certain adaptation solving that problem, such as jealousy. The products of these heuristics are adaptationist hypotheses, i.e., hypotheses that explain traits as adaptations. Indeed, both reverse engineering and adaptive thinking are adaptationist strategies that ascribe to natural selection the primary role in the explanation of traits.

Evolutionary Psychologists have firmly defended the adequacy of these heuristics. They argue that, despite being a young field, Evolutionary Psychology has proven itself capable of suggesting plenty of evolutionary hypotheses and advancing our understanding of

human psychology and that the adaptationist methodology has a crucial role in making this possible. For instance, Cosmides and Tooby claim that the adaptationist methodology “can free cognitive scientists from the blinders of intuition and folk psychology, allowing them to construct experiments capable of detecting complex mechanisms they otherwise would not have thought to test for” (Cosmides & Tooby 1994, 41). Focusing on the evolutionary function of traits is seen as “an indispensable methodological tool, crucial to the future development of cognitive psychology” (Cosmides & Tooby 1994, 43). According to several Evolutionary Psychologists, this is exactly the purpose of Evolutionary Psychology. As Schmitt and Pilcher (2004) explain, Psychologists take the view that “identifying all the psychological adaptations that make up human nature – the adaptationist program of humanity – is what evolutionary psychology should be all about” (2004, 644). Recently, a group of Evolutionary Psychologists has published a guide for conducting research in Evolutionary Psychology (Lewis et al. 2017). In this guide, Psychologists confirm their confidence in adaptive thinking and reverse engineering. Indeed, these heuristics are suggested as the appropriate, reliable methodological tools for generating rigorous hypotheses (Lewis et al. 2017, 355).

Do these discovery heuristics make Evolutionary Psychology fruitful? As mentioned, fruitful programs use *reliable* heuristics, namely heuristics that avoid defective reasoning. Let’s analyse the reasoning involved in reverse engineering and adaptive thinking.

Reverse engineering involves observing a current trait and inferring from it the adaptive problem that the trait possibly solved in EEA. Adaptive thinking involves first hypothesizing an adaptive problem faced in EEA, and then inferring from that the specific trait that was evolved to solve the problem. Both these heuristics assume “a strong relationship between adaptive forces and the resulting organism” (Sterelny & Griffiths 1999, 242). This assumption of optimality may hinder the formulation of well-designed hypotheses. Formulating well-designed hypotheses on the evolutionary nature of psychological traits is not an elementary work. The evolution of traits involves several aspects that should be meticulously considered in the formulation of hypotheses. For instance, developmental and environmental factors do play a role in the evolution of traits by constraining the action of natural selection. Because of such a strong optimality assumption, it is not clear how these heuristics can account for these factors. As Paul Griffiths (2001) explains, the inferences allowed by these heuristics (i.e. inferring the adaptive problem from the solution and inferring the solution from the adaptive problem) are then unreliable. These heuristics may be too imprecise to constitute reliable tools for formulating evolutionary hypotheses. This means that they may constitute an obstacle to the development of the research program. Lewis and colleagues have recently argued that Psychologists do consider various evolutionary factors in the formulation of hypotheses, such as by-products (Lewis et al. 2017, 362). However, whether the consideration of these factors is widespread as claimed should be assessed by means of a careful analysis of the hypotheses suggested by Psychologists in their papers. In

order to be accurate, the consideration of different evolutionary factors should be present in any evolutionary study of psychological traits.

Recently, Edouard Machery (2011) has argued that although the adaptationist heuristics may produce inaccurate hypotheses, this does not mean that they systematically do that. He claims that the adaptationist strategy may actually be beneficial, since it reduces the class of possible hypotheses. Why is this beneficial? Machery does not clarify the benefits of reducing the class of hypotheses. However, there are at least two reasons why this may be beneficial. First, it makes scientists' work easier. Having a restricted number of hypotheses to analyse makes it easier to find out what is the best hypothesis among them. Second, this strategy may be beneficial because it admits only the hypotheses that are worth considering. It rules out bad hypotheses, namely hypotheses that are likely to be false. One of the benefits of this strategy is then preventing proliferation of hypotheses.

In fact, these heuristics prevent the formulation of many hypotheses. Is this a virtue of the program? As mentioned above, these heuristics face the problem of focusing exclusively on adaptationist hypotheses and ignoring other kinds of evolutionary hypotheses, such as the ones involving by-products. This means that they ignore plausible evolutionary hypotheses that need to be considered to reach an exhaustive analysis of the evolution of traits. This is not a desirable reduction of hypotheses. Moreover, these heuristics may actually multiply hypotheses by suggesting too many adaptationist hypotheses, also the ones that are likely to be false but are considered because of their adaptationist nature. The adaptationist strategies, then, do not prevent the proliferation of hypotheses. They actually *encourage* it. By preventing the formulation of many legitimate evolutionary hypotheses, these strategies undermine the possibility to improve and extend our understanding of human psychology and this is a problem for the fruitfulness of Evolutionary Psychology.

The discovery heuristics used in Evolutionary Psychology have what Gould calls "virtuosity in invention" (Gould 1978, 530). They are powerful tools enabling scientists to imagine scenarios explaining the existence of a trait and its evolutionary function. Still, because of the unreliable inferences they allow, their ability to formulate well-designed hypotheses is undermined. Their power to develop Evolutionary Psychology is then limited.

3.3 Is Evolutionary Psychology fruitful?

The previous sections pinpoint the problems of the research question and heuristics used in Evolutionary Psychology. These problems constrain the power of the program to extend and improve the understanding of human psychology and, then, they compromise its fruitfulness. Still, there is room for improvement. The adaptationist question is not necessarily fruitless. For instance, in evolutionary medicine this question is used as a fruitful tool to improve our understanding of medical conditions and, in some cases, our ability to cure diseases (Nesse & Stearns 2008). The adaptationist question is just one of the evolutionary questions used to

guide research in evolutionary medicine. Indeed, as Randolph Nesse and Stephen Stearns explain, research in evolutionary medicine involves both the adaptationist question “How has mechanism M given a selective advantage to individuals?” and the general evolutionary question “What is the phylogeny of this mechanism?” (Nesse & Stearns 2008, 30). Evolutionary Psychology, then, might improve its fruitfulness by employing a more pluralistic methodology, namely involving both adaptationist and non-adaptationist questions. The evolutionary factors question suggested by Lloyd seems to be a good option to improve Evolutionary Psychology’s fruitfulness. Moreover, adaptationist heuristics are not doomed to failure in any context. By analysing research in evolutionary biology, Sara Green noticed that while methodological adaptationism may be a productive heuristic for some specific tasks, such as exploring the connections between causal capacities and selective pressures, it performs poorly for other tasks, such as studying drift (Green 2014, 493). The same might be true for Evolutionary Psychology. It is then necessary to understand in which cases adaptationist heuristics may be able to develop the program and in which cases they hinder the development.

Psychologists often praise the fruitfulness of their program. They frequently provide lists of the successes of Evolutionary Psychology, such as new discoveries and theories (e.g. Buss & Reeve 2003). It is an open question whether all the items on these lists do constitute desirable development, i.e., extensions of our understanding of human psychology. Indeed, many of the discoveries and theories of the program have been criticized as being the products of poor scientific research, even the ones that are usually presented as triumphs of the program, such as the Sexual Strategies Theory suggested by Buss and Schmitt (1993). Evolutionary Psychology is a very creative program. Still, its research question and heuristics compromise its fruitfulness.

4. Conclusion

I have argued that the traditional accounts of fruitfulness are too vague to provide a strategy to assess the fruitfulness of research programs. I have suggested a new approach, which is focused on analysing the set of tools used by scientists to extend and improve the content of a program. Specifically, this approach focuses on examining research questions and discovery heuristics.

Despite having an explication of fruitfulness and a strategy to assess it, scientists’ assessments of programs may still clash with each other. For instance, scientists may disagree on which research questions are more reliable. Nonetheless, having precise criteria to evaluate fruitfulness provides the basis for unambiguous and useful debates and makes it possible to avoid several disputes over the merits of programs: to be fruitful, programs have to use suitably detailed research questions and reliable discovery heuristics. The explication

I suggest in this chapter constitutes a starting point for a serious discussion of the actual relevance of fruitfulness in the assessment of research programs. However, the analysis of fruitfulness can be developed further by examining other tools, besides research questions and heuristics, that can be used to improve and extend the content of programs, such as the methodologies to test hypotheses and predictions.

As discussed in the introduction, philosophers of science make a distinction between epistemic and cognitive values (e.g. Laudan 2004; see section 1.2 of Chapter 1). While epistemic values are truth-indicative features of programs and inform us about their truth value, cognitive values do not tell anything about the truth value of programs, but are valuable for pragmatic reasons, such as making research easier and faster. Is fruitfulness an epistemic or cognitive value?

There is disagreement over the desirability of fruitfulness in the philosophical literature, i.e., whether fruitfulness is desirable for epistemic or pragmatic reasons. Some philosophers argue that fruitfulness is an epistemic value (McMullin 1983), while others claim it is a cognitive value (Kuhn 1977; Douglas 2009).

The case study discussed in this chapter shows that fruitfulness is a virtue of scientific research programs, which have to be able to develop and improve our understanding or knowledge of the world. Moreover, fruitfulness may facilitate the fulfilment of other values. Using research questions that permit an exhaustive analysis of phenomena and heuristics that enable scientists to formulate well-designed hypotheses promotes the fulfilment of values like predictive power, scope, and explanatory power. However, in the next chapter, I will discuss what determines the desirability of values and I will argue that in order to understand the ground for the desirability of values, contextual factors (i.e. aspects of the context in which a value is used to assess theories) should be considered. Various factors may contribute to the fruitfulness of programs, such as local contingencies and the skills of scientists. For instance, the funds available to programs and scientists' competence can contribute to determine the development of programs. Considering these aspects may be important for understanding the type of information provided by the value and the reasons of its desirability.

Chapter 3

The desirability of values in theory appraisal

A context-sensitive approach

1. Introduction

David Buss and Hudson Kern Reeve have recently argued that

“By critical scientific standards – theoretical cogency, predictive accuracy, interdisciplinary consistency, and empirical harvest – modern evolutionary psychology fares well compared with alternatives.” (Buss & Reeve 2003, 848)

Buss and Reeve use ‘scientific standard’ to refer to what is usually called ‘value’ by philosophers of science.²² In this quote, they make two points. First, Evolutionary Psychology satisfies some important values, such as predictive accuracy. Second, Evolutionary Psychology compares well in terms of scientific standards to its alternative research programs.²³

²² See section 1.5 of the first chapter for a discussion over the terminology used in the philosophical literature.

²³ Buss and Reeve also claim that the alternative to Evolutionary Psychology suggested in Lickliter and Honeycutt 2003 (namely Developmental Dynamics, which is the main target of Buss and Reeve’s critique in that paper) is an “obscure alternative” that does not perform better than Evolutionary Psychology in terms of scientific standards, such as predictive accuracy, testability, and parsimony (Buss & Reeve 2003, 851).

To motivate their appeal to these specific values, Buss and Reeve claim that features such as theoretical cogency are critical standards holding in all scientific disciplines (Buss & Reeve 2003, 849). As discussed in the introduction, this claim is widely shared among philosophers of science, who have argued that values like predictive accuracy and consistency are features that are often relevant to the assessment and comparison of alternative theories or research programs (e.g. Kuhn 1977; McMullin 1983). The idea is that scientists can legitimately appeal to certain features to assess scientific theories and research programs because these are valuable features of theories and programs. They are valuable because it is desirable for theories and programs to have these features (McMullin 1983, 5; see also section 1.1 of the introduction).

The desirability of values is usually explained by making a distinction between epistemic and cognitive values (e.g., Laudan 2004; see also section 1.2 of the introduction). Epistemic values are truth-indicative features of scientific theories, namely features that are desirable because informative about the truth value or reliability of scientific theories; cognitive values are instead desirable because they make those theories easier or faster to work with, thereby providing scientists with pragmatic reasons for preferring a theory over an alternative (Douglas 2009, 93). The distinction, then, is based on the different reasons provided by epistemic and cognitive values to prefer a theory over its alternatives.

A question that needs to be addressed is: what is it that makes a value epistemically and/or pragmatically desirable? In other words, in virtue of what can a value like fruitfulness or simplicity provide scientists with good reasons to believe that a theory is true or worthy of being used further? This chapter is dedicated to explore this question.

The chapter is structured as follows. In the second section, I present my case study, namely the Evolutionary Theory of Stalking suggested in Duntley & Buss (2012) and I use explanatory power, external consistency, and fruitfulness to analyse it. Moreover, I focus on the value of explanatory power to compare the Evolutionary Theory of Stalking to another theory of stalking, namely the integrative account proposed by Davis and colleagues (2012). In the third section, I argue for three main theses about values and their importance in theory assessment.

First, I will argue for the *situatedness of theory appraisal*. My claim is that certain contextual factors determine whether a given value can provide information relevant to the assessment of theories. I will argue that the assessment of theories is done at a specific time in history and in a specific cultural context, involving scientific, methodological, political, and social factors. These contextual factors may determine what values can tell scientists about those theories, by facilitating or obstructing *scientists' understanding* of whether and to what extent a theory satisfies a value.

Second, I will argue for the *contextual desirability of values*. My claim here is that certain contextual factors determine whether a value is desirable for epistemic or cognitive reasons. I will then argue that the desirability of a given value like fruitfulness or explanatory

power may vary across contexts and historical periods.

Third, I will argue for the *interdependence between values*. Whether a theory satisfies a certain value and what this tells scientists about the theory often depend on whether the theory satisfies other values. The epistemic or cognitive importance of a value may vary, depending on how and whether the theory satisfies other relevant values.

While little attention has been paid to the influence of contextual factors on the assessment of theories in the existing philosophical literature, I conclude the third section by discussing the importance of my three claims for clarifying the epistemic/cognitive desirability of values and for providing an exhaustive account of the legitimate roles that values play in science. In the fourth section, I will make some final remarks.

2. Explanatory power, external consistency, and fruitfulness of the Evolutionary Theory of Stalking

Let's start by defining the values that will be used in the assessment of the Evolutionary Theory of Stalking.

The value of *explanatory power* concerns – unsurprisingly – the power of a theory to explain phenomena. There is an ongoing debate on what *explaining* phenomena means and philosophers have suggested various theories of explanation. For instance, according to the deductive-nomological theory (e.g. Hempel 1965), something counts as an explanation of something else if it provides a sound deductive argument, in which what has to be explained deductively follows from a set of true premises that includes at least a law of nature. According to the unificationist theory, instead, to explain is “to fit the phenomena into a unified picture” (Kitcher 1989, 500), by deriving “descriptions of many phenomena, using the same pattern of derivation again and again” (Kitcher 1989, 432). Other theories of explanation are the causal-mechanical model (e.g. Salmon 1984; Woodward 2003; Bechtel & Richardson 1993) and the pragmatic theory of explanation defended by Bas van Fraassen (van Fraassen 1980).

A point of contention is whether there is only one model able to account for the concept of explanation or rather a plurality of concepts of explanation, and each model is “correct in capturing some salient aspect of explanatory practice” (Colombo 2017, 507).²⁴ In this dissertation, I will take explanations to be answers to some why-, how-, when-, or where-questions. An answer to a request for explanation is a good one to the extent it either unifies apparently scattered pockets of knowledge about the phenomenon of interest, or can be used

²⁴ Research in psychology has revealed that people give explanations by appealing to unifying regularities (e.g. Williams and Lombrozo 2010) as well as by referring to causal mechanisms (e.g. Walker, Lombrozo, Legare, & Gopnik, 2014). The pluralistic view, then, seems to better account for the way human beings use the concept of explanation.

to address counterfactual questions about what would happen if certain features of the phenomenon of interest were different. A theory is able to provide explanations (i.e. it has explanatory power) if it can satisfactorily answer a sufficient number of relevant questions about why and how certain phenomena happen and it is able to make those phenomena unsurprising and expected (Peirce 1935; Shupbach & Sprenger 2011).

Some argue that explanatory power is an epistemic value. For instance, Dana Tulodziecki (2011) argues that a close look at the history of science reveals that explanatory power has played an epistemic role in science, where explanatory considerations have often promoted the attainment of knowledge.²⁵ Others, instead, claim explanatory power is a cognitive value since it provides no information over the truth value of a theory (e.g. van Fraassen 1980, 71; Douglas 2009, 107; see also Douglas 2013, 800). That explanatory power does not always play an epistemic role is illustrated by just-so stories, such as the fables written by Rudyard Kipling to explain the origin of the distinctive features of some animals (e.g., the trunk of the elephants; see Kipling 1902/1989), which can provide us with explanations that make certain phenomena unsurprising, but are just inventive stories short on evidence, and with little power to provide answers to counterfactual questions about those phenomena.

The second value used in the appraisal of the case study is *external consistency*. The value is included in virtually all the lists of values suggested by philosophers (e.g. Kuhn 1977; McMullin 1983; Douglas 2009; see also Longino 1995), and it concerns the consistency of a theory with other scientific theories. A theory is consistent with other theories if its claims do not contradict the claims made in those other theories; that is, if the theory under assessment and the theories with which it is evaluated provide pictures of phenomena that do not conflict with each other. The theories that scientists regard as immediately relevant to assess the external consistency of a theory may vary across contexts. It may depend on the field a theory belongs to, or the research program in which scientists assessing the theory work (see McMullin 2009 for a categorization of the different levels in which the external consistency of a theory may be evaluated).

Douglas (2013) includes external consistency in the group of epistemic values. However, she points out that the epistemic importance of external consistency may vary across contexts: whether external consistency is truth indicative depends on the truth or reliability of the background theories (Douglas 2013, 798; see also Steel 2010, 20). If we

²⁵ Tulodziecki focuses on the explanatory power of the theory of the communication of cholera suggested by the British physician John Snow (Tulodziecki 2011, 306). The theory performed better than its competitors in providing answers to questions on how and why certain phenomena happened (e.g. why doctors visiting patients and performing the post mortem during the epidemics did not contract the disease) (Tulodziecki 2011, 312). These explanatory considerations, Tulodziecki argues, were important to understand the accuracy of the theory and played a crucial role in the comparison of alternative theories accounting for the communication of cholera (2011, 307).

have reasons to doubt the truth or reliability of the background theories, external consistency has little or no epistemic importance, since it would not inform us over the truth value of the theory.

The third value used in the case study is *fruitfulness*. The value concerns the ability of a theory to extend its content and suggest new avenues to extend our understanding of phenomena (e.g. Kuhn 1977; Douglas 2009; see also Chapter 2). In the second chapter, I have suggested we should focus on research questions and discovery heuristics if we want to assess the fruitfulness of scientific theories and research programs (see also Ivani 2019).

In the literature, there is disagreement on why fruitfulness may be a desirable feature of scientific theories. As discussed in the previous chapter, McMullin sees fruitfulness, and in particular the virtue he calls proven-fertility, as an epistemic value (McMullin 1976, 400). Douglas, instead, argues that fruitfulness is a cognitive value: it is valuable for pragmatic reasons, since it may make research easier and faster (Douglas 2009, 93). By providing new lines of research, a fruitful theory allows for new avenues to test the theory, i.e., it provides scientists with more opportunities to test the theory. However, providing new lines of research is not *in itself* a sign that the theory is true (Douglas 2009, 93; Douglas 2013, 800).²⁶

Let's now turn to the case study. In the next sections, I will introduce stalking as studied in the legal and psychological literature. I will then present the Evolutionary Theory of Stalking and I will analyse its explanatory power, external consistency, and fruitfulness. Explanatory power will also be employed to compare the theory to another account of stalking suggested in the psychological literature, namely the integrative account of stalking (Davis et al. 2012).

2.1 Stalking behaviours and sex differences

Some people may find being rejected by a partner (or potential partner) particularly hard to accept. Some may even engage in persistent attempts to establish or re-establish the romantic relationship or to control the partner in order to avoid the rejection. These attempts may involve non-consensual communication, unwanted physical proximity and attention targeting an individual that does not reciprocate the desire of a romantic relationship.

Psychologists call the “ongoing and unwanted pursuit of romantic relationships between individuals [who are either] not currently involved with each other or who have broken up with each other” *persistent pursuit* (Davis et al. 2012, 329) and the “repeated and unwanted pursuit and invasion of one's sense of physical or symbolic privacy by another person [...] who desires and/or presumes an intimate relationship” *obsessive relational*

²⁶ It should be noted that McMullin and Douglas provide different explications of fruitfulness. While for McMullin, proven-fertility concerns the (past) success of a theory in growing and transforming itself to accommodate anomalies (McMullin 1976, 400; see also Chapter 2 of this dissertation), in Douglas' view fruitfulness is the ability of a theory to provide new lines of research, theoretical implications, and predictions (Douglas 2009, 93).

intrusion (Cupach & Spitzberg 1998, 234-5). When these attempts cause fear or other negative reactions in the victim, or a “reasonable person (or jury) would view [the behaviours] as fearful or threatening” (Spitzberg & Cupach 2007, 66), persistent pursuit and obsessive relational intrusion are treated as cases of *stalking* (Lyndon et al. 2012, 300; see also Caretti et al. 2011, 7).²⁷

The psychological and legal literature presented in this case study consider the following elements to identify stalking behaviours. First, stalking involves a course of conduct: the perpetrator of the behaviours has to persistently attempt to establish contact or intimacy with the victim. Second, these behaviours are unwanted by the victim. Examples are being followed or spied and receiving calls, messages, and gifts. Finally, these behaviours may cause some negative reactions in the victim or in a reasonable person (e.g. distress, serious alarm or fear).

However, the importance of the negative reactions triggered by these behaviours to identify stalking cases varies across studies. In fact, the way stalking is defined, studied, and prosecuted varies across countries. For instance, in the US legal system only the cases in which the victim *experiences* fear are treated as stalking (e.g. Lyndon et al. 2012; Smith et al. 2018). The British legal system, instead, does not make this requirement, and treats as stalking also cases in which the victims do not experience fear (Lyndon et al. 2012, 300). In other legal systems, the cases identified as stalking are the ones in which the perpetrator of the unwanted attention has the *purpose* of threatening and frightening the victim (e.g. Article 285b of the Dutch Criminal Code). Which definition is used, and whether certain behaviours are treated as a case of stalking may depend on cultural assumptions, such as what people in a community normally see as acceptable behaviour in romantic relationships and courtship. Such assumptions may vary across societies and historical periods.²⁸

Depending on the definition used, findings on stalking frequency may vary. The percentages of stalking victims usually rise when the fear requirement is not included (see Lyndon et al. 2012, 300 for a discussion of this issue). However, despite some variation, studies on stalking frequency report consistent data on sex differences. They all report that female victims outnumber male victims. For instance, the National Intimate Partner and Sexual Violence survey reveals that, in the US, the 16.0% of the female population and the 5.8% of the male population have been victims of stalking at some point in their lifetime (Smith et al. 2018, 5). Cupach and Spitzberg have conducted a meta-analysis of 51 studies

²⁷ Stalking behaviours may also have other aims than maintaining, establishing or re-establishing a romantic relationship, such as revenge and controlling co-workers. This case study focuses on the explanations of stalking behaviours aimed at continuing, establishing, and re-establishing romantic relationships.

²⁸ Several studies have been conducted to understand how stalking is perceived across cultures. See Lippman (2018) for a study on how culture influences the way people perceive stalking behaviours and Jagessar and Sheridan (2004) for a study on the differences and similarities of stalking perception across cultures.

on stalking and found that 75% of the victims were females (Cupach & Spitzberg 2004, 48). The study has also revealed sex differences in perpetrators of stalking behaviours (22% females, 77 % males, Cupach & Spitzberg 2004, 48; see also Lyndon et al. 2012, 300).²⁹

Moreover, studies reveal females and males differ in the tactics used. Males are reported to engage in highly aggressive persistent pursuit (e.g. pursuit involving physical violence) more often than females do (Davis et al. 2012). Females, instead, “engage in less serious forms of persistent pursuit such as following, showing up uninvited, and persistent telephoning, texting, and emailing” (Davis et al. 2012, 332). It should be noted that the majority of the studies on stalking have been conducted in the U.S., UK, and Australia. Spitzberg and Cupach report that US samples constitute almost the 70% of the available studies (Spitzberg & Cupach 2007, 66). The generalizability of the claims on frequency and sex differences is then limited.

However, a variety of samples have been studied, such as clinical patients, general population, and college students. The research based on samples of college students reveals sex similarities in frequency of persistent pursuit, while the other samples reveal the aforementioned sex differences. Both female and male college students engage in persistent pursuit (Shorey et al. 2015, 939). However, sex differences have been observed in the strategies used by female and male college students, with males leaving unwanted items for the partner to find more often compared to females (Shorey et al. 2015, 938).

The existing literature on stalking behaviours is extensive. Several theories have been suggested to explain why some individuals engage in these behaviours and the reasons for the observed sex differences in frequency and tactics (e.g. Davis et al. 2012; Meloy 1999). One of these theories is the Evolutionary Theory of Stalking suggested by Joshua Duntley and David Buss (Duntley & Buss 2012), which will be presented and analysed in the next sections. It should be noted that Duntley and Buss use ‘stalking’ to refer to behaviours causing fear in the victims as well as behaviours not inducing fear in the victims. The reason is that Duntley and Buss aim at explaining *why* and *how* some individuals engage in certain patterns of behaviours (such as persistently engaging in behaviours aimed at establishing or re-establishing a relationship with an individual). These behaviours may or may not induce fear in the victims (since victims may or may not be aware of being spied and followed, Duntley & Buss 2012, 312). But, according to Duntley and Buss, these behaviours can all be explained by referring to the same (evolutionary) origin. From here on, I will use the term ‘stalking’ to refer to persistent behaviours aimed at establishing, re-establishing or maintaining a romantic relationship with an individual who does not reciprocate such a desire, behaviours that may (or not) induce fear in the victims.

²⁹ While in the National Intimate Partner and Sexual Violence Survey Smith and colleagues use a definition of stalking involving the fear requirement (i.e. only the cases in which victims *experience* fear are treated as stalking), the meta-analysis developed by Cupach and Spitzberg involves studies using a fear-free definition as well as studies employing a definition in which fear is included.

2.2 The Evolutionary Theory of Stalking

Duntley and Buss explain stalking mechanisms in terms of adaptations, that is, in terms of evolved psychological mechanisms that have been selected due to the fitness advantage they provided (see section 2.1 of Chapter 1).³⁰ They say:

“We propose that stalking can be beneficial by contributing to the solution of recurrent ancestral problems affecting reproductive success. Because mating is central to reproductive success, we theorize that stalking evolved as one among a menu of strategies designed to solve ancestrally recurrent problems relevant to human mating.” (Duntley & Buss 2012, 314)

Duntley and Buss' idea is the following. Our ancestors have faced some recurrent problems that threatened their reproductive success. Stalking mechanisms helped our ancestors to solve these adaptive problems. Individuals equipped with stalking mechanisms were relatively more successful (in some contexts) in mating than their competitor. Duntley and Buss specify that stalking is just one of the strategies that may successfully enhance reproductive success. They argue that it may be “a last-ditch desperation tactic when other avenues for solving the problem of mating are perceived to be closed off” (Duntley & Buss 2012, 312).

Among the adaptive problems that stalking strategies may have helped to solve, Duntley and Buss include:

“acquiring a new mate, guarding an existing mate to prevent defection, fending off potential mate poachers, poaching away someone else's mate, strategically interfering with competitors for mates, reacquiring an ex-mate, sexual exploitation and predation, guarding female mates and kin to prevent them from being sexually exploited.” (Duntley & Buss 2012, 316)³¹

According to Duntley and Buss, humans have evolved a wide array of stalking mechanisms, and each one of them was specifically successful in solving one of the adaptive problems in terms of reproductive success in EEA.

An example of stalking mechanism is controlling the everyday life of the partner by monitoring their movements and isolating and monopolizing their time. The mechanism can be activated by different triggers in men and women. For instance, men may engage in stalking behaviours when mated to fertile, young, and physically attractive women or when

³⁰ Duntley and Buss claim that some design features of stalking, such as the use of weapons, may “have been exapted from hunting adaptations and warfare adaptations” (Duntley & Buss 2012, 315).

³¹ Human beings are not the only species facing these problems and stalking tactics are used by other species as well, such as non-human primates (Duntley & Buss 2012, 317). However, the theory presented in this case study specifically focuses on human beings' stalking strategies resulting from these problems.

partners seem interested in rivals with better economic resources (Duntley & Buss 2012, 317). Women, instead, are hypothesized to engage in stalking behaviours when their partners display signs of emotional involvement with potential rivals (Duntley & Buss 2012, 317-318).

Duntley and Buss argue that selection has favoured this mechanism to solve the adaptive problem of the defection of the partner (2012, 317). Losing a partner has costs in terms of reproductive success: it impedes sexual access and the possibility to reproduce with that partner. Engaging in stalking behaviours like the ones just mentioned may make it possible to avoid these costs, maintaining the relationship, and ensure reproductive success. Likewise, when our ancestors faced the problem of reacquiring an ex-mate, stalking strategies could be helpful for regaining sexual access, and thus increasing the individual's reproductive success (Duntley & Buss 2012, 318-9). For instance, Duntley and Buss claim that offering gifts could serve the function of showing the victim how committed the individual engaged in these tactics is to start a relationship (2012, 316).

In Duntley and Buss' view, both females and males have evolved stalking mechanisms, since both females and males have faced adaptive problems that stalking mechanisms may have helped to solve.³² Moreover, they hypothesize that females and males have evolved stalking mechanisms in response to the sex-specific problems they faced and that different factors trigger stalking mechanisms in females and males (some examples will be provided in the next paragraphs). However, it is important to notice that this is not to say that females and males completely differ in their stalking strategies. In Duntley and Buss' view, females and males have evolved similar stalking strategies when facing similar problems, or when the same strategy could solve the different adaptive problems faced by females and males.

The aforementioned sex differences in frequency and tactics of stalking presented in the legal and psychological literature can then be explained as the result of the different evolutionary pressures females and males experienced in EEA. Let's have a closer look at how Duntley and Buss explain the sex differences in stalking mechanisms.

One of the assumptions at the basis of their theory is that women and men experience significantly different amounts of parental investment. Parental investment is defined as “any investment by the parent in an individual offspring that increases the offspring's chance of surviving (and hence reproductive success) at the cost of the parent's ability to invest in other offspring” (Trivers 1972, 139). It is argued that females experience a more considerable amount of parental investment compared to males because of their obligatory investment. Indeed, unlike males, females get pregnant and breastfeed, which are necessary to the birth and survival of the offspring. Males' investment, instead, can amount to the only contribution

³² In what follows, I follow Duntley and Buss' terminology and use the terms *females*, *women*, *males*, and *men* as synonyms.

of their sperm for the conception of the offspring (Buss & Schmitt 1993, 206).³³

The differences in parental investment, Duntley and Buss claim, are “a clear source of evolutionarily recurrent conflict between the sexes [...] and central to understanding the evolution of stalking” (2012, 313). The reason is that the asymmetry in parental investment poses different adaptive problems to females and males and it contributes to determine which sexual strategy can maximize their reproductive success.

Because of their considerable parental investment, females would face the problems of finding protection and resources for themselves and for their offspring. Seeking and keeping long-term monogamous relationships with a committed partner would be an effective strategy to solve these problems (Buss & Schmitt 1993, 206). Instead, since the parental investment of males can be minimal in terms of time- and energy-consumption, males would face the problem of finding as many females as possible to impregnate in order to maximize their reproductive success. Short-term strategies (brief affairs with no commitment to a romantic relationship, such as one-night stands) would then help solve males' problems in terms of reproduction (Buss & Schmitt 1993, 206).³⁴

Evolutionary Psychologists claim that the studies they have conducted corroborate these hypotheses. Cross-cultural studies report that women and men differ in sexual desires: they differ in the number of partners desired and in the type of relationships they want³⁵ (e.g. Buss & Schmitt 1993; Schmitt et al. 2003; Schmitt et al. 2012; see also Chapter 4 of this dissertation). Men are reported to desire more partners than women do and to prefer short-term strategies, whereas women are reported to prefer committed partners and long-term strategies (Buss & Schmitt 1993, 204).

When it comes to stalking tactics, Duntley and Buss hypothesize that females and males have evolved some sex-specific tactics because of the different adaptive problems faced. To illustrate, let's consider the hypothesis concerning the use of high-risk stalking strategies (e.g. violence and highly threatening tactics). Duntley and Buss argue that males, because of their limited amount of parental investment, can benefit from using these strategies more than females. The reason is the following.

Reproductive variance is greater among males than among females. While most

³³ Several factors are involved in the calculation of parental investment of women and men, which are discussed in Chapter 4.

³⁴ This is an idealized model. That is, given the evolutionary pressures and the amount of parental investment for women and men, these are the strategies that can better enhance the reproductive success of women and men. However, studies report that women and men seek different types of sexual strategies, i.e., women and men look for both short- and long-term partners (e.g. Buss & Schmitt 1993). See Chapter 4 for a discussion of women's and men's preferences for different sexual strategies.

³⁵ It should be noted that many of the studies collect data on sexual desires by means of surveys, in which participants are asked which type of relationship and number of partners they desire. Participants' responses might then differ from the actual type of relationship and number of partners they have in real life.

fertile women succeed in reproducing during their lifespan and they all have a similar number of children (to a maximum of 12), men can significantly differ in their reproductive success, with some men failing to reproduce and others having dozens of children (Buss 1999/2008, 100; see also Bateman 1948 on reproductive variance). Unlike males, females have then little reasons to engage in risky strategies to acquire a mate. Indeed, males who take risks, such as engaging in violent tactics targeting competitors and potential mates, can prevail over their competitors and have sexual access to women. Males who are able to impregnate multiple females simultaneously maximize their reproductive success and deprive their competitors of mating opportunities (Duntley & Buss 2012, 314). Duntley and Buss therefore hypothesize that males have evolved violent stalking strategies and that they may engage in violent tactics more frequently than females in present-day societies.

Moreover, Buss and Duntley argue that because of the sex-specific problems posed by parental investment, different factors may *trigger* different stalking tactics in females and males. For instance, they hypothesize that some stalking tactics, such as mate-guarding strategies, are used by females when suspecting *emotional* infidelity, namely when suspecting the partner being romantically involved with someone else (Duntley & Buss 2012, 322). Having a partner being romantically involved with someone else may mean that he will end the relationship and start a new one with the other partner.³⁶ The defection of the partner may lead one to lose resources and protection necessary for women's reproductive success and the survival of the offspring. By guarding partners' behaviours, then, women can avoid the defection of the partner from the long-term relationship and avoid the reproductive costs involved in the defection.

Instead, stalking strategies would be activated in males when suspecting *sexual* infidelity, namely when suspecting the partner is sexually involved with another individual (Duntley & Buss 2012, 322). While women are always certain that they are genetically related to their offspring, males are not (Duntley & Buss 2012, 314). Controlling the behaviours of the partner may then help men to avoid sustaining somebody else's offspring and increasing somebody else's reproductive success rather than their own.

Let's now turn to the appraisal of the Evolutionary Theory of Stalking. Does this theory satisfy the values of explanatory power, external consistency, and fruitfulness? I will start the analysis by considering the value of explanatory power.

2.3 Explanatory power of the Evolutionary Theory of Stalking

I have defined explanatory power as the ability of a theory to satisfactorily answer questions about why and how certain phenomena happen. A theory is able to explain a given phenomenon when it makes the phenomenon unsurprising and expected. Now, what's the

³⁶ It should be noted that these statements on what triggers stalking behaviours in women and men focus on *heterosexual* relationships.

degree of explanatory power of the Evolutionary Theory of Stalking?

The Evolutionary Theory of Stalking provides us with an answer to why some individuals but not others engage in stalking behaviours. Certain individuals engage in stalking behaviours because our ancestors evolved mechanisms for stalking as successful solutions to the mating problems faced in EEA and those individuals encounter situations that activate those mechanisms. Explaining stalking behaviours as evolutionary solutions to mating problems makes relatively unsurprising that individuals engage in stalking when facing mating problems.

For instance, consider the problem of regaining an ex-partner. Regaining an ex-partner means reacquiring sexual access to that partner and having the possibility to enhance one's own reproductive success. The theory makes unsurprising that some individuals engage in stalking when they try to reacquire an ex-mate. Studies report that a significant amount of stalkers (half of the total number, according to De Smet et al. 2013 and Spitzberg & Cupach 2007) are ex-partners of the victims.³⁷

The Evolutionary Theory of Stalking seems also able to explain the aforementioned sex differences in stalking behaviours. One of its fundamental assumptions is that females and males have faced sex-specific problems in EEA because of differences in parental investment. Because of these differences, Duntley and Buss hypothesize that females and males have evolved different stalking strategies and that stalking strategies can be activated by different factors in females and males (e.g. types of jealousy).

To illustrate the ability of the theory to explain sex differences, take the use of aggressive stalking tactics (e.g. murder and physical and sexual violence). As mentioned, Duntley and Buss argue that, because of reproductive variance, aggressive stalking tactics were more useful to men than to women in solving adaptive problems (e.g. gaining sexual access to a potential partner). They then hypothesize that men recur to aggressive stalking tactics more often than women do.

This hypothesis can explain the observed sex differences in the use of aggressive tactics by women and men. Indeed, studies reveal men engage in highly aggressive stalking mechanisms more frequently than women do. For instance, Tjaden and Thoennes (1998) report that male-to-female intimate partner violence is more frequent than female-to-male intimate partner: 25% of female participants and 8% of male participants reported rape or physical assaults perpetrated by a former or current partner (Tjaden & Thoennes 1998, 2). Hence, if it is true that, because of evolutionary reasons, aggressive tactics were more useful to men than to women, the theory makes it unsurprising that men engage in these tactics more frequently than women do.

³⁷ More precisely, the theory is able to explain why we find that one of the motivations underlying stalking behaviours is regaining access to a partner. However, it does not necessarily explain the fact that a *considerable number* of stalkers engage in these behaviours because of this motivation.

The Evolutionary Theory of Stalking seems therefore to have some degree of explanatory power. It provides answers to questions regarding how and why certain phenomena happen and makes the phenomena to be explained unsurprising.

However, it is not clear whether the answers the Evolutionary Theory of Stalking provides are the *right* ones, namely whether these constitute accurate explanations of why and how people engage in stalking behaviours. Understanding whether the theory really explains the phenomenon of stalking and the strength of its explanations might be a complicated task and the reason is the following.

The evidence able to show that the theory provides us with accurate explanations of how stalking behaviours evolved and of the ways present-day humans engage in such behaviours should be able to demonstrate the *evolutionary history* of stalking mechanisms. That is, it should be able to justify the historical claim that the evolution of stalking behaviours happened in EEA. Duntley and Buss claim that many stalking mechanisms are adaptations. Evidence, then, should be able to support the hypothesis that these mechanisms were able to solve actual adaptive problems and that they actually solved these adaptive problems in EEA, that selection of the trait actually occurred, and that the trait is heritable (Brandon 1990; Richardson 2007). These are evidential requirements that any theory aimed at explaining the evolutionary history of a trait has to meet (see Richardson 2007). A theory providing an account on the evolution of a trait unable to meet these evidential requirements may be seen as a speculative story.

To date, this evidence is lacking. As mentioned, the theory is anchored in the research conducted in evolutionary biology and Evolutionary Psychology on the parental investment and sexual desires of women and men (e.g. Buss & Schmitt 1993). Moreover, the theory offers explanations able to make the data on frequency, motivations, and sex differences of stalking behaviours unsurprising. However, Duntley and Buss do not provide evidence specifically able to show that the evolution of stalking went the way they hypothesize.

For instance, no evidence aimed at corroborating the hypothesis that the stalking mechanism of mate guarding was evolved as a solution to prevent defection of the partner (e.g. the heritability of the trait) has been provided. Duntley and Buss are aware of this limitation and admit that they have just “presented some limited empirical evidence” able to support their theory (for instance the studies on jealousy as one of the most frequent motives to engage in stalking behaviours, Duntley & Buss 2012, 324). They then argue that “the power and utility of the theory must be evaluated in the future” when the relevant evidence in neuroscience, anthropology, and animal behaviour will be collected (for instance they mention the research on the neurobiology of stalking conducted by Meloy & Fisher 2005; Duntley & Buss 2012, 324). According to Duntley and Buss, the hypotheses derived from their theory are testable and falsifiable (2012, 321). For instance, they mention that one of the hypotheses of the theory is that stalking strategies are cross-culturally universal and

“finding that stalking is not adopted in other cultures to address the adaptive problems described in this paper would refute our theory” (Duntley & Buss 2012, 321). However, according to Duntley and Buss, what we know about human mating, such as the function of jealousy and the sexual desires of women and men, makes the explanations provided by the Evolutionary Theory of Stalking plausible (Duntley & Buss 2012, 318).

A question that may be asked is whether the corroboration of hypotheses is relevant to ascribe explanatory power to a theory. Granted, the relevant evidence to corroborate the hypotheses of the theory is lacking, but does this really undermine the explanatory power of the theory?

Explanatory power was defined as the ability of a theory to satisfactorily answer questions concerning why and how certain phenomena happen, by making these phenomena unsurprising. Corroboration might be irrelevant to assess this ability of the theory. By appealing to parental investment and adaptive problems, the Evolutionary Theory of Stalking is able to answer questions on why and how human beings might possibly engage in stalking behaviours and why we observe sex differences in frequency and tactics of stalking.

In order to clarify the explanatory power of the theory, some questions that need to be addressed are: What does it take for a theory to *satisfactorily* answer explanatory questions? Should satisfactory answers to explanatory questions always be corroborated by evidence?

A way to deal with these issues is differentiating between types of explanatory power that a theory may have. Specifically, the explanatory power of scientific theory may be actual or potential. The Evolutionary Theory of Stalking seems to lack the former and satisfy the latter. The difference lies in the type of explanations that the theory is able to provide. While a theory with actual explanatory power provides us with how-actually explanations, a theory with potential explanatory power provides how-possibly explanations.

An explanation qualifies as a how-actually explanation if we have substantial evidence supporting the explanation (Brandon 1990; Durrant & Haig 2001). Evidence, then, makes us confident that the explanation accurately accounts for the phenomena. How-possibly explanations, instead, can be defined as follows:

“An explanation is a how-possibly explanation if and only if it 1) lacks adequate empirical support, but 2) still satisfies other explanatory virtues.” (Resnik 1991, 143)

How-possibly explanations do not have evidence showing that they accurately account for certain phenomena. Nonetheless, they are valuable because they offer potential explanations for which we have little in evidence in favour, but also limited evidence against. These explanations offer interesting lines of research, which are worthy of being pursued because they satisfy some other important virtues (e.g. simplicity and external consistency). For instance, Resnik argues that how-possibly explanations are valuable because they can be

fruitful, with heuristic power, functioning as basis for formulating novel hypotheses (Resnik 1991, 143; see also Durrant & Haig 2001, 360).

The Evolutionary Theory of Stalking does not provide us with how-actually explanations of the evolution of stalking. However, it provides us with how-possibly explanations which might be worthy of being pursued because consistent with some fundamental claims of Evolutionary Psychology, such as the claims about different sexual desires of women and men.

Moreover, the how-possibly explanations provided by the theory may be valuable as a fruitful basis for the development of novel hypotheses that may enhance our knowledge of stalking behaviours. For instance, hypothesizing that because of different adaptive problems, different factors may trigger stalking mechanisms in females and males (e.g. types of jealousy) may offer new lines of research to understand the factors that specifically lead women and men to engage in stalking behaviours.

Whether the theory provides fruitful how-possibly explanations will be discussed in the next section. However, before turning to the assessment of the fruitfulness and external consistency of the Evolutionary Theory of Stalking, we should consider another question about its explanatory power: does the Evolutionary Theory of Stalking perform better than other theories of stalking in terms of explanatory power? Take the account of stalking suggested by Davis and colleagues (2012) as an example. Davis and colleagues propose to integrate three psychological theories, namely coercive control theory, relational goal pursuit, and attachment theory, to explain why some individuals engage in stalking behaviours and why there are some sex differences in stalking. Stalking behaviours can be explained as coercive control, i.e., behaviours aimed at controlling other individuals' lives. These behaviours arise in individuals who lack self-regulation and perceive an ex-partner as the exclusive means to satisfy some personal goals (such as having a satisfying romantic relationship) and experience intense distress caused by the separation from this individual (Davis et al. 2012, 333).

These individuals experience intense distress when rejected because of the tendency in attachment they developed in infancy. The bonding with caregivers that individuals experience during their childhood has an important evolutionary function in infancy (i.e. it secures the survival of the infants) and it affects the way individuals experience relationships in adulthood. According to Attachment Theory, “securely attached infants derive security from accessible and responsive caregivers, whereas insecurely attached infants experience rejection, abandonment, or indifference from primary caregivers” (Spitzberg & Cupach 2007, 77). Individuals who have developed insecure attachment in their infancy may experience difficulties in handling unreciprocated desires of a romantic relationship and engage in stalking behaviours.

Davis and colleagues argue the cultural context plays a major role in determining how women and men engage in stalking behaviours. In this view, the aforementioned sex

differences in frequency and strategies of stalking would be the product of the cultural context: culturally determined gender roles and expectations determine which behaviour is appropriate for men and women and facilitate the spread of certain behaviours among the members of each sex. Davis and colleagues argue that the cultural context³⁸ favours “male dominance and men's role as the sexual aggressor” (2012, 332). Men are pushed towards authoritative and dominant roles, while women are pushed to be coy and submissive. These differences in gender roles and expectations would then have a central role in determining different frequency in highly aggressive stalking strategies in women and men.

It should be noted that the integrative account does not rule out the evolutionary component from stalking mechanisms, and it does not explain the sex differences as the mere product of culture. The idea is that human psychology is determined by various factors, such as the evolutionary factor of attachment to caregivers, and, given the expectations and gender roles of the cultural context, some individuals end up engaging in stalking behaviours. Insecure attachment may lead women and men to behave in certain ways, and these behaviours may be amplified, encouraged or inhibited by the gender roles and expectations of the cultural context. The Evolutionary Theory of Stalking and the integrative account, then, should not be seen as accounts appealing to completely different, mutually inconsistent factors in their explanations of stalking behaviours.

The integrative account of stalking provides explanations of why some individuals engage in stalking behaviours. Davis and colleagues argue that the theories integrated in their account have already proved to be promising in the explanation of stalking behaviours, in particular in the explanation of stalking behaviours in clinical cases (i.e. individuals with attachment disorders; Spitzberg & Cupach 2007, 77). Moreover, the integrative account makes the sex differences in tactics and frequency reported in the legal and psychological literature unsurprising. If it is true that the cultural context plays a major role in pushing women and men towards different behaviours – and we may have good reasons to think that this is true (see for instance the available data on the influence of gender norms on sexual lives in Sanchez et al. 2005 and managerial styles in Holmes & Schnurr 2005) – it is not surprising to observe the aforementioned sex differences in frequency and strategies of stalking.³⁹

Now, does the Evolutionary Theory of Stalking perform better than the integrative account in terms of explanatory power? To compare the explanatory power of the two accounts, I suggest we consider the ways they explain some of the sex differences observed in stalking behaviours aimed at regaining an ex-partner. Studies report that a huge portion of

³⁸ As mentioned, the studies on stalking behaviours have mainly used samples from U.S., U.K., and Australia. (Spitzberg & Cupach 2007, 66). These countries, then, constitute the cultural context Davis and colleagues refer to.

³⁹ This means that we might find that women and men engage in stalking behaviours differently depending on the gender norms accepted in the societies they live in.

stalkers are ex-partners of the victim (e.g. De Smet et al. 2013; Spitzberg & Cupach 2007). Explaining stalking behaviours aimed at regaining an ex-partner is then one of the major goals that a theory of stalking can achieve.

One of the basic statements of the Evolutionary Theory of Stalking is that the different amount of parental investment had a role in the evolution of sex-specific sexual strategies and stalking mechanisms. Specifically, because of the limited amount of obligatory parental investment, men are hypothesized to prefer short-term strategies and numerous sexual partners. Men are then hypothesized to have evolved stalking strategies that were successful in solving their adaptive problems, such as being able to mate with more than one partner.

The explanatory power of this hypothesis, namely whether this hypothesis can explain the data on frequency of stalking behaviours aimed at regaining an ex-partner in women and men, may be called into question. If it is true that men are evolutionarily driven to prefer short-term strategies, it is surprising to find that men engage in stalking behaviours to regain access to a previous partner when they can maximize their reproductive success by looking for a new short-term relationship with another partner. More precisely, it is surprising to find that men do so more often than women do, who are supposed to be more interested in long-term strategies and, according to Evolutionary Psychologists, experience higher costs in terms of reproduction and survival when losing a partner compared to men. Indeed, Evolutionary Psychologists argue that women, by losing a long-term committed partner, would lose access to resources (e.g. food), protection (for themselves and the offspring), opportunities for learning, and status (Buss & Schmitt 1993, 222).

Duntley and Buss argue that being rejected lowers one's own value as a mate (i.e. after a rejection, the individual might be perceived as an undesirable mate by possible future partners) and this may undermine the chance of finding new mates and enhancing one's own reproductive success (Duntley & Buss 2012, 319). This explanation provides a plausible account of stalking behaviours aimed at regaining a partner, but it is not clear whether this can explain the sex differences reported in the legal and psychological literature. This explanation assumes that men experience higher costs in terms of mate value when rejected compared to women and it is not clear why this should be the case.

A reason for this claim might be the difference in reproductive success among women and men. Evolutionary Psychologists argue that while women have similar reproductive success (i.e. similar number of children), men's reproductive success may significantly vary, with some men failing to reproduce and others having dozens of children. Being rejected by a partner, then, may seriously undermine the chances that a man has to mate and reproduce. Stalking mechanisms might then have evolved to prevent or avoid men's reproductive failure.

The claim that there are significant sex differences in reproductive variance is based on the studies conducted by geneticist Angus John Bateman, which reported significant sex

differences in reproductive success in *Drosophila melanogaster* (Bateman 1948). This difference in reproductive variance is one of the central assumptions of Evolutionary Psychology, underlying its theories on human mating. However, recent experimental studies have failed to replicate Bateman's results of the significant sex differences in reproductive success (see Gowaty et al. 2012) and the soundness of this fundamental assumption might need to be reconsidered.

Now, consider the integrative account. In this account, the data on men's stalking behaviours are not surprising: the cultural context (e.g. gender norms on how men and women should be and how they should conduct themselves in romantic/sexual contexts) of the countries in which the stalking studies have been conducted pushes men to engage in aggressive, dominant stalking behaviours to regain access to an ex-mate, whereas women do not experience the same cultural pressure. When it comes to explaining these sex differences in stalking behaviours, the integrative account seems to perform better than the Evolutionary Theory of Stalking.

However, both the integrative account and Duntley and Buss' theory face the following problem. Whether the available data provide an accurate picture of sex differences in stalking behaviours is open to question. Indeed, it is not clear whether they report the actual frequency and types of stalking tactics used by women. The frequency with which women engage in persistent pursuit and stalking behaviours may have been underestimated (Davis et al. 2012, 337). Research has paid little attention to some of women's behaviours, such as flirting, seduction, and using one's own physical attractiveness, which are used as persistent pursuit strategies. Moreover, women's aggressivity may have been underestimated. For instance, little is known about the aggressiveness of women's behaviours aimed at damaging the reputation of the desired partner (or ex-partner). Another point is that men might report to be victim of stalking less frequently than women do. The psychological literature offers mixed results. Some studies report that women and men do not differ in which behaviours they categorize as cases of stalking (Sheridan et al. 2002), while others have found some differences, revealing that men are less likely to qualify various cases of persistent pursuit as stalking (Yanowitz 2006; see also Lyndon et al. 2012 for a discussion of this issue).

Gathering more data on the frequency and aggressivity of women's stalking tactics and on how men experience being victim of stalking behaviours is necessary for achieving a better understanding of the phenomenon. Having these points clarified will enable us to evaluate the actual explanatory power of the two accounts. For instance, if studies reveal that women and men engage in aggressive stalking tactics at similar rates, the two accounts might need to revise some of their statements. For instance, the Evolutionary Theory of Stalking may need to revise the way it explains the use of aggressive tactics (i.e. by appealing to differences in reproductive variance).

2.4 External consistency and fruitfulness of the Evolutionary Theory of Stalking

Although the Evolutionary Theory of Stalking may face some problems in terms of explanatory power, the theory could still be evaluated as a good, promising theory if it satisfies other relevant values.

When discussing the virtues of the Evolutionary Theory of Stalking, Duntley and Buss emphasize its consistency with Sexual Strategies Theory (Buss & Schmitt 1993). Sexual Strategies Theory hypothesizes the aforementioned sex preferences in sexual strategies on the basis of the different parental investment experienced by women and men. Sexual Strategies Theory is frequently presented as one of the triumphs of Evolutionary Psychology, showing the predictive power and fruitfulness of the program (e.g. Buss & Reeve 2003). The consistency of the Evolutionary Theory of Stalking with this theory, then, may be seen as good news and of primary importance for its appraisal.

Several studies have been conducted to test the hypotheses of the Sexual Strategies Theory, producing evidence about the hypothesized sex differences in sexual strategies (e.g. Schmitt et al. 2003; Schmitt et al. 2012). For instance, one of the predictions of the theory is that, because of the different strategies able to maximize reproductive success, men are more open to casual sex compared to women. The prediction has been tested in cross-cultural studies by collecting data on how much time women and men wait before consenting to have sex with someone. Studies reveal sex differences in permissive sexuality: the likelihood of intercourse after knowing someone for a short time (e.g. 1 week) is much higher among men (positive about the possibility of consenting to sex after one week) than among women (highly unlikely to have sex after knowing someone for just 1 week) (Buss & Schmitt 1993, 211; see also Surbey & Conohan 2000).

Sexual Strategies Theory has also been widely criticized. Critics have emphasized two problems. First, the fundamental assumptions of the theory have been called into question. Specifically, the way the parental investment of women and men is calculated is unconvincing, since men's amount of investment for the survival of the offspring is not adequately considered (e.g. Hubbard 1992). For instance, the evaluation of men's parental investment used by Evolutionary Psychologists ignore the energy involved in male-male competition and in the protection of the offspring (see the fourth chapter for an extensive discussion of the evaluation of parental investment of men and women). If these costs were considered, critics argue, it would turn out that men's parental investment plays a more substantive role in the survival of the offspring than previously claimed (Hrdy 1986). This means that the hypothesis concerning the preferences of sexual strategies of women and men based on the significantly different parental investments of women and men should be reconsidered.

Second, it is argued that Evolutionary Psychologists have cherry picked data to confirm their theory. The theory hypothesizes that men and women differ considerably in sexual desires and preferences for sexual strategies, but, according to critics, data do not show

that women and men significantly differ in their sexual desires. For instance, Andrew Smiler argues that existing data reveal substantial similarities in sexual strategies of women and men, for instance in the number of partners desired by women and men (Smiler 2011, 604). Evolutionary Psychologists should take into account the data that have been deliberately ignored and reconsider their confidence over the predictive accuracy of the theory (Smiler 2011, 604; see also Chapter 4 of this dissertation).

The Evolutionary Theory of Stalking is consistent with Sexual Strategies Theory. Taken together, these two theories provide a consistent, harmonious portrayal of the strategies evolved by humans to enhance their reproductive success. However, the criticisms just discussed indicate that whether Sexual Strategies Theory is able to account for human mating is open to question, and the consistency of the hypotheses and predictions of the Evolutionary Theory of Stalking with this theory might then be bad news.

The consistency with Sexual Strategies Theory is essential to the Evolutionary Theory of Stalking, since the former provides the guidelines for the formulation of the hypotheses and predictions of the latter on stalking strategies. Hence, if Sexual Strategies Theory is an inaccurate, biased account of human mating, this consistency may call into question the soundness of the Evolutionary Theory of Stalking and it may then be seen as an (epistemically) undesirable quality of the Evolutionary Theory of Stalking. Contrary to what is argued by Duntley and Buss, the consistency of the Evolutionary Theory of Stalking with Sexual Strategies Theory does not speak in favour of the truth or reliability of the former.

What about the fruitfulness of the theory? As discussed in the second chapter, the value of fruitfulness concerns the ability of a theory to extend its content and expand our knowledge. I have suggested assessing fruitfulness by considering the properties of the tools used to extend the content of a theory in an actual research context, such as research questions and heuristics. The analysis of the fruitfulness of the Evolutionary Theory of Stalking will focus on research questions.

The central questions of the theory are evolutionary questions. Specifically, Duntley and Buss ask what the evolutionary functions of stalking behaviours might be. For instance, Duntley and Buss focus on how stalking mechanisms can solve the adaptive problems of regaining sexual access to an ex-partner and interfering with competitors to have sexual access to a possible partner (Duntley & Buss 2012, 316).

Do these types of questions have the power to make the theory fruitful? The way Evolutionary Psychologists use adaptationist questions may in fact undermine the fruitfulness of the theory. Specifically, when the adaptationist questions are the *only* questions asked and adaptationist answers are the only answers seen as legitimate, the ability of the theory to enhance our understanding of a phenomenon may be limited (see the discussion developed in section 3.1 of Chapter 2).

Understanding whether using an adaptationist methodology (i.e. focusing only on adaptationist questions) can extend our knowledge of stalking involves understanding

whether non-adaptationist evolutionary questions (e.g. examining how stalking strategies were exapted from other mechanisms or whether some of these strategies may be by-products) are of secondary importance or may be ignored. However, as it was pointed out in the second chapter, the evolution of a trait involves various factors (e.g. environmental and developmental constraints). Hence, exploring non-adaptationist evolutionary questions is actually relevant to understand stalking.

In the second chapter, I have explored the importance of asking non-adaptationist questions in order to make Evolutionary Psychology's theories fruitful. In this section, I will focus the discussion on fruitfulness on its relations with the values of external consistency and explanatory power.

In the discussion of the external consistency of the Evolutionary Theory of Stalking, I have claimed that the consistency of the theory with Sexual Strategies Theory is fundamental for the theory of stalking. Several evolutionary questions used to develop the theory of stalking are based on the sex differences in sexual desires posited by Sexual Strategies Theory.

I have already emphasized how the consistency of the Evolutionary Theory of Stalking with Sexual Strategies Theory may be epistemically undesirable, but this consistency may affect the fruitfulness of the Evolutionary Theory of Stalking too. Taking Sexual Strategies Theory as fundamental basis to formulate questions about stalking might constrain and limit the fruitfulness of the Evolutionary Theory of Stalking. Assuming that men and women considerably differ with respect to their preferred sexual strategies, and assuming that this has a fundamental role in determining how and whether they engage in stalking behaviours, might be a weak starting point to formulate questions on stalking behaviours. This is because the problems and biases of Sexual Strategies Theory might be transmitted to the Evolutionary Theory of Stalking. If it is true that consistency with Sexual Strategies Theory is undesirable, i.e., if it is not epistemically desirable to be consistent with a theory for which many criticisms have been raised, asking questions formulated on the basis of this theory may negatively constrain the ability of the Evolutionary Theory of Stalking to extend our understanding of stalking.

In the discussion of the explanatory power of the theory, I have argued that the Evolutionary Theory of Stalking might be able to provide fruitful how-possibly explanations. That is, the theory might be able to provide explanations (temporarily) lacking empirical support but still valuable because fruitful. However, if critics are right that the statements concerning the different sexual strategies are unwarranted and evidence does not support them, the external consistency of the theory with Sexual Strategies Theory might then question the ability of the theory to provide fruitful how-possibly explanations.

On a more general level, the analysis of the external consistency and fruitfulness of the Evolutionary Theory of Stalking may shed light on the fruitfulness of Evolutionary Psychology as a research program. I have defined a fruitful program as one capable of

extending its content by producing novel theories and hypotheses. Building up on Sexual Strategies Theory, Duntley and Buss have proposed the Evolutionary Theory of Stalking, which provides novel hypotheses on human psychology and its evolution. This development of the content of the program may question the fruitfulness of the program. If it is true that Sexual Strategies Theory is a weak starting point to formulate novel hypotheses, this development of the content of the program could be seen as an epistemically unsuccessful development. Putting the issue in Lakatosian terms, the fact that from the basic theories of Evolutionary Psychology (e.g. Sexual Strategies Theory), theories that we have good reasons to think as false or inaccurate are developed, may lead us to think that Evolutionary psychology is a *degenerating* research program (Lakatos 1970; see also section 2.1 of Chapter 1). In other words, Evolutionary Psychology would be a program that produces novel hypotheses that - once the relevant evidence will be collected - will be systematically falsified.

The fact that the Evolutionary Theory of Stalking may turn out to be an undesirable development or extension of the content of Evolutionary Psychology does not allow us to conclude that Evolutionary Psychology is a *fruitless* program. Other theories suggested by Evolutionary Psychologists may be seen as epistemically successful developments and extensions of the program. However, if Sexual Strategies Theory constitutes the basis of many of the extensions of the program and we have good reasons for thinking that it is a problematic theory, whether the theory is the best way to make the program fruitful should be considered carefully.

3. Discussion: what the case study tells us about values

In the previous section, three values – explanatory power, external consistency, and fruitfulness – have been used to analyse the Evolutionary Theory of Stalking (Duntley & Buss 2012). Explanatory power has also been used to compare the theory to the integrative account of stalking (Davis et al. 2012).

What can we say about the nature and role of values in the evaluation of scientific theories in light of the case study?

First, the case study shows that understanding whether a theory satisfies a value may be hampered by local contingencies. Understanding whether the two theories of stalking have explanatory power is obstructed by the unavailability of evidence. The evidence relevant to assess the evolutionary explanations provided by the Evolutionary Theory of Stalking is lacking and the reliability of the available data on sex differences in stalking behaviours is open to question. These factors make it difficult to understand whether the two theories are actually able to explain why some individuals engage in stalking behaviours and the ways women and men engage in such behaviours. Were the relevant evidence available, we would

have clearer ideas on which theory performs better in terms of explanatory power.⁴⁰

The case study then shows that contextual factors, such as the availability of relevant evidence, sometimes have an impact on the evaluation of whether and how a theory satisfies a value. The valuable features of a theory may be ignored and underestimated (or overestimated), because the context in which the theory is evaluated does not allow for an accurate assessment of whether and how the theory satisfies certain values. Contextual factors may then determine whether the ascription of a value to a theory (such as “Theory T has explanatory power”) is justified, i.e., whether we have good reasons to claim so.

The case study then reveals that the evaluation of whether and how a theory satisfies a value is *historically and culturally situated*. I call this thesis *the situatedness of theory appraisal*. The evaluation of scientific theories is always done at a specific time in history, in which technologies and evidence relevant to the assessment of theories may be lacking, and in a specific cultural context, involving scientific, methodological, political, and social factors. This means that, although a theory may actually satisfy a certain value, depending on the local contingencies of a historical and cultural context, we might be unable to appreciate this.

In the case study presented in this chapter, the local contingency hampering the understanding of the explanatory power of the theories of stalking was the unavailability of relevant evidence, a factor concerning the scientific practice. However, other types of contextual factors may be relevant in determining the way a scientific theory is assessed.

Consider the way the theory of female orgasm as by-product was initially received by the scientific community, example that was presented in the introduction of this dissertation (Symons 1979; Lloyd 2005). Although the theory of female orgasm as by-product had great explanatory power (the theory was able to explain the evolutionary origin of female orgasm and the data on frequency and modalities of orgasm in women) and performed better than its alternatives in terms of explanatory power (see Lloyd 2005 for a discussion of alternative theories), scientists were unable to acknowledge its explanatory power because of certain contextual factors.

The explanatory power of the theory was ignored for a long time because the theory did not provide an adaptationist explanation of female orgasm. At that point in the history of science, adaptationism dominated scientific research and by-product explanations were seen as having secondary importance compared to adaptationist explanations. More precisely, as Lloyd 2015 explains, non-adaptationist explanations were considered only when adaptationist explanations failed to be corroborated by evidence (Lloyd 2015, 349; see also the introduction and the second chapter of this dissertation). Moreover, the theory explaining female orgasm as by-product was also criticized by some feminist scholars because it was

⁴⁰ This means that, when contextual factors like the lack of evidence obstruct an accurate assessment of theories, we may want to suspend our judgment on whether a theory satisfies a certain value until the relevant evidence is produced.

perceived as a sexist account of female sexuality. Specifically, since the theory provides a non-adaptationist explanation of female orgasm, the theory was seen as ascribing female sexuality secondary importance in the evolutionary history of human sexuality (Lloyd 2001).

The appreciation of the explanatory power of the by-product theory of female orgasm, then, was obstructed by both science-related and society-related factors, namely the methodological assumptions of adaptationism endorsed by the scientific community and the feminist agenda endorsed by some members of the scientific community.⁴¹

Let's now turn to the second point. The case study shows that contextual factors may determine whether and why values are desirable in relation to a given theory. I call this thesis *the contextual desirability of values*. Take external consistency as an example. External consistency is included in virtually all the lists suggested in the philosophical literature (see section 1.1 of the introduction). The reasons for its desirability or whether it is desirable at all may vary across contexts, depending on the accuracy or reliability of the relevant background theories (see Douglas 2013; Steel 2010). If we have good reasons to think that the background theories are accurate or reliable, the consistency with those theories may be seen as a sign that our theory is accurate or reliable (or not obviously false).

The case study shows that external consistency – in that context – is *undesirable*. More precisely, the consistency of the Evolutionary Theory of Stalking with Sexual Strategies Theory is epistemically undesirable because of the problems of the latter. Indeed, if it is true that Sexual Strategies Theory faces the difficulties discussed in the previous section (namely problems concerning its basic assumptions and empirical support), being consistent with such a theory is not indicative of the truth or reliability of the Evolutionary Theory of Stalking. If it is true that Sexual Strategies Theory is an inaccurate portrayal of human mating, taking this theory as a basis for formulating evolutionary explanations of stalking may actually obstruct the attainment of knowledge.

In that context, external consistency is not an epistemic value. However, it should be noted that this does not show that external consistency is not a *valuable* feature of scientific theories. Ideally, external consistency is a valuable feature of scientific theories because of the aims ascribed to science, such as providing a consistent picture of the world. Moreover, it does not show that external consistency cannot be truth indicative: the consistency of a theory with true or reliable theories provides us with reasons to think that the theory might be true or reliable. What the case study shows is that *in this case* external consistency is not truth indicative and that a theory may satisfy an important value, such as external consistency, but whether this is *desirable* depends on whether the consistency involves true or reliable theories or not.

⁴¹ This is not to say that the social and political values of a cultural context always impede an accurate assessment of theories. As I will discuss in the fourth chapter, social factors as feminist values may play an epistemic role in the assessment of scientific theories, i.e., they may enhance the acquisition of knowledge.

Moreover, although the consistency of the Evolutionary Theory of Stalking with Sexual Strategies Theory is not truth indicative, the value of external consistency – in that context – is still valuable. The reason is that external consistency is able to provide information relevant to the assessment of the Evolutionary Theory of Stalking, namely information over possible problems of the theory due to its consistency with Sexual Strategies Theory. Values, then, are those features of scientific theories that, although they might be epistemically undesirable in a specific context, they are still valuable because able to provide relevant information over theories, also in a “negative” way.

The third and last thesis discussed in this section is the thesis of *the interrelationships between values*. Our case study shows that values can be interrelated with each other. In the case study developed in this chapter, considerations about the external consistency of the Evolutionary Theory of Stalking with Sexual Strategies Theory were relevant to the assessment of the fruitfulness and explanatory power of the theory. Indeed, the consistency of the theory with Sexual Strategies Theory may question its fruitfulness and explanatory power, i.e., its ability to enhance our understanding of stalking and provide accurate explanations.

Understanding whether and how a theory satisfies a certain value (e.g. external consistency) may be relevant to understanding whether and how the theory satisfies other values (e.g. fruitfulness and explanatory power). The relations among values may then play an important role in theory appraisal, by determining the kinds of reasons values can provide to the assessment of theories and determining the weight that these values can have in the assessment.

Overall, this study strengthens the idea that epistemic and cognitive values constitute important tools to understand whether a theory is true, accurate, reliable or worthy of being developed further. Moreover, it highlights that whether a value can actually provide reasons to believe a theory true or reliable, involves considering that the assessment is done in a specific historical and cultural context as well as the relations that values may have with each other.

To date, the literature in philosophy of science has paid little attention to the influence of contextual factors on theory appraisal (e.g. Douglas 2013; see also Steel 2010). However, the three theses discussed in this section should be integrated in the accounts aimed at clarifying the roles that values can play in the assessment of scientific theories. In particular, they are relevant if we think that the roles that a value can legitimately play in theory appraisal depends on whether it is epistemically or pragmatically desirable (e.g. Douglas 2009).

4. Conclusion

This chapter was dedicated to exploring what makes values able to provide epistemic and/or cognitive reasons to assess scientific theories and prefer a certain theory over its alternatives. To develop this analysis, I have presented a case study, namely the Evolutionary Theory of Stalking, and I have used three values to analyse it. The case study shows that contextual factors (namely elements of the historical and cultural context in which a theory is assessed) are among the factors determining the desirability of values.

Before turning to the analysis of non-cognitive values in Evolutionary Psychology, a clarification. The case study presented in this chapter focuses on research in psychology. The conclusions on the desirability of values and on the relations among values may then be non generalizable to other fields.

However, this should not be seen as a limitation of the argument developed in this chapter. My analysis shows that, in order to understand the effective importance and desirability of values, we should consider the local, specific factors involved in the assessment of theories. This implies that understanding the roles that values can have in science necessarily involves analysing case studies from different fields.

As mentioned in the second section, the way stalking is studied and prosecuted varies across countries. Partially different definitions are used in legal systems and psychologists have employed different notions of stalking to study the phenomenon. Cultural assumptions about romantic relationships, courtship, and gender roles may play a role in determining these differences. Although the phenomenon of stalking has not been considered in the philosophical literature, it constitutes an interesting case study deserving attention. The case study developed in this chapter focused on cognitive values, but stalking may constitute an interesting case study for future research on how objects of study are conceived (namely what should count as stalking behaviour) and the ways non-cognitive values, cultural expectations, and gender roles may influence this process.

The next chapter will focus on the influence of non-cognitive values in the assessment of one of the theories analysed in this case study, namely Sexual Strategies Theory (Buss & Schmitt 1993). Specifically, the next chapter will focus on the beneficial roles that non-cognitive values – in particular feminist values – have played in the assessment of the theory.

Chapter 4

Feminist critiques of Evolutionary Psychology

The cognitive role of non-cognitive values⁴²

1. Introduction

The role of non-cognitive values in the assessment of scientific theories has been hotly debated in the philosophical literature. According to the value-free ideal, only cognitive values⁴³ play a legitimate role in theory appraisal, whereas non-cognitive values should be excluded from these phases of scientific research (see section 1 of Chapter 1). The reason is that, unlike cognitive values, non-cognitive values cannot help with the attainment of knowledge. When non-cognitive values influence theory appraisal, advocates of the value-free ideal argue, they undermine the reliability of scientific reasoning and the epistemic authority of science (e.g. Haack 1993).

However, the value-free ideal has been widely criticized as a descriptively inaccurate and normatively undesirable portrayal of science. Critics of the value-free ideal have argued that the history of science and scientific practice show that non-cognitive values can play relevant, legitimate roles in the assessment of scientific theories. For instance,

⁴² This chapter is based on joint work with Jan Sprenger.

⁴³ In this chapter, I will use ‘cognitive values’ to refer to both cognitive and epistemic values.

Longino (1990) emphasizes the role of non-cognitive values in evidential reasoning. Rudner (1953), Hempel (1965), and Douglas (2000) have argued that non-cognitive values play a legitimate role in theory appraisal by influencing the assessment of cases on inductive risk.

In this chapter, I aim at showing that non-cognitive values can play a cognitive role in the assessment of scientific theories, since they can help scientists acquire new knowledge of reality, just like cognitive values do. To do so, I will present a case study concerning the epistemically beneficial influence of feminist values on theory appraisal.

Feminist scholars are interested in detecting and fighting forms of oppression of women. When it comes to science, feminists are especially committed to detect the ways in which scientific practice neglects women as an object of research and denies them an active role in the production of knowledge. In the last decades, a considerable feminist literature has emerged in the fields of sociobiology, evolutionary biology, and psychology (for an overview see Gowaty 1997). The case study presented in this chapter will highlight the role of feminist values in combating bias in the account of human sexuality in Evolutionary Psychology (Buss & Schmitt 1993).

The chapter is structured as follows. In the second section, I expound the case study. I first introduce the evolutionary account of human sexuality. I then present the feminist critique of this account and I discuss the cognitive role that feminist values play in this case study. In the third section, I argue that, since non-cognitive values can play a cognitive role in theory appraisal, the traditional cognitive/non-cognitive distinction of values needs to be reconceived. However, I will also argue that not any non-cognitive values can play the cognitive role that feminist values play in the case study and that a distinction between different types of non-cognitive values is needed. Section 5 summarizes the findings and concludes.

2. Human sexual behaviour and Evolutionary Psychology

Evolutionary Psychologists have conducted several studies to clarify the evolutionary origins of human mating, revealing that humans have evolved a variety of traits, such as preferences for partners with certain features (Buss et al. 1990; Buss, Shackelford, Kirkpatrick, & Larson, 2001), tactics of mate poaching (Schmitt & Buss 2001; Duntley & Buss 2012), and emotional responses to first-time intercourse (Haselton & Buss 2001). Research on human sexuality in Evolutionary Psychology has been tremendously prolific and extensively presented and discussed in newspaper articles, radio, and television (Cassidy 2005; see also section 2.1 of Chapter 1). As the philosopher of science David Buller reports, you might read about Evolutionary Psychology's research on sexual behaviours on the covers of women's and men's magazines "during almost every wait in the supermarket checkout line" (Buller 2005, 3). Because of the astonishing amount of studies conducted, their alleged success in

corroborating the predictions of the account, and their popularity outside academia, Evolutionary Psychologists usually present their account of human sexuality as one of the triumphs of Evolutionary Psychology (e.g. Buss & Reeve 2003).

In this section, I analyse the influence of non-cognitive values on the account of human sexuality provided by Evolutionary Psychology. In particular, I focus on Sexual Strategies Theory (Buss & Schmitt 1993), and the theories on which it is based, namely Sexual Selection Theory and Parental Investment Theory. Let's start with the most time-honored theory: Sexual Selection Theory.

Charles Darwin observed that in several species sexes differ in respect to many features, such as size, colours, and behaviour. Not all of these features can be explained by natural selection, namely as features evolved because solving adaptive problems of survival. For instance, Darwin was puzzled by the flashy train of peacocks: how can a peacock survive and defend himself with that iridescent train that makes him more “like a dandy than a warrior”? (Darwin 1871, 43). In order to explain these anomalies, Darwin formulated the theory of Sexual Selection: individuals look for partners with good genes, i.e., partners who can provide them with a healthy offspring able to survive. The “enigmatic” features, such as the peacock's train, indicate the good genes of individuals and help them to find a partner and spread their genes.

Two mechanisms are especially important for sexual selection. First, the members of one sex are involved in *intrasexual competition*, also called ‘male-male competition’, namely a competition to have sexual access to the other sex. Second, the other sex is involved in the *intersexual selection*, also called ‘female choice’. Individuals choose to mate with individuals that display certain features, i.e., features that signal the good genes of the potential partner. Evolutionary Psychologists think that while men are involved in the intrasexual competition, women make the selection.

Sexual Selection Theory, by itself, does not explain this difference in roles between sexes. An explanation was provided by Parental Investment Theory. The theory was first formulated by Robert Trivers, who defined parental investment as

“any investment by the parent in an individual offspring that increases the offspring's chance of surviving (and hence reproductive success) at the cost of the parent's ability to invest in other offspring.” (Trivers 1972, 139)

The difference in roles would be a result of the different amount of parental investment experienced by sexes. While one sex invests less in offspring and is involved in intrasexual competition, the other sex invests more and makes the intersexual selection. Trivers' argument hinges on anisogamy, namely sexual reproduction involving gametes that differ in respect to quantity and size. This asymmetry in gametes would determine the amount of parental investment of sexes. The sex that produces few large gametes invests more in

offspring. The reason is that these gametes are rare and precious and, in order to avoid wasting them, the individuals producing these cells choose carefully the partner that can assure good use of them. The sex that generates many small, cheap gametes, on the other hand, invests less in offspring and is involved in the intrasexual competition. When it comes to humans, women produce large gametes, i.e., ova, and men generate small and cheap gametes, i.e., sperm. Hence, traditionally, evolutionary biologists and psychologists claim that women have a greater parental investment than men. Indeed, women are involved in an ‘obligatory’ investment that is indispensable for the birth and survival of the offspring, including nine months of pregnancy and breastfeeding.

On the basis of Sexual Selection Theory and Parental Investment Theory, Evolutionary Psychologists David Buss and David Schmitt propose Sexual Strategies Theory to explain human sexual behaviour (Buss & Schmitt 1993). Buss and Schmitt claim that humans have evolved two main sexual strategies, namely long-term and short-term strategies. While a long-term strategy aims at long, monogamous relationships, a short-term strategy does not lead to a real commitment to a romantic relationship. Buss and Schmitt argue that

“men and women have faced different mating problems over human evolutionary history, at least in some delimited domains, and therefore the principles that govern the mating of women and men are predicted to be different in these domains.” (Buss & Schmitt 1993, 205)

Buss and Schmitt hypothesize that the different amounts of parental investment pose different adaptive problems to women and men and determine which sexual strategy they are more willing to pursue. While men would be more inclined to short-term mating than women, women would look for long-term relationship more often than men.⁴⁴ By means of short-term strategies, men would solve their adaptive problem posed by the big quantity of cheap gametes. That is to say, since men’s reproductive success is constrained by the number of women they can inseminate, short-term mating would allow them to use their big quantity of gametes and spread their genes. By pursuing long-term strategies, women would solve the adaptive problems posed by their obligatory investment. Long-term mating would allow them to find partners that provide protection and resources that are necessary for the survival of the offspring.

Evolutionary Psychologists have produced an astonishing amount of studies and empirical evidence that would support the predictions of Sexual Strategies Theory on the existence of sex differences in mating strategies (e.g. Buss et al 1990; Schmitt et al. 2012). Manifold studies have detected sex differences in the desired number of partners (Schmitt et al. 2003). Furthermore, several studies show that while women value partners with resources

⁴⁴ These claims are statistical. Buss and Schmitt concede that women and men use both strategies (see for instance Buss & Schmitt 1993).

and high status more than men (Gottschall et al. 2003), men prefer partners with traits that signal fertility (e.g. Buunk et al. 2001; Confer et al. 2010).

Sexual Strategies Theory assigns an essential role to the amount of parental investment in the evolution of sexual behaviours. However, this is done on the basis of contentious and sometimes outright questionable assumptions. In the next section, I demonstrate how feminist values have helped to uncover and challenge these assumptions, and how they improved the reliability and objectivity of evolutionary accounts of human mating.

2.1 Feminist critique of the evolutionary account of human mating

The account of human sexuality suggested in Evolutionary Psychology and biology has been widely criticized. Many of these criticisms have been raised by scholars who explicitly embrace feminist values (e.g. Hubbard 1990; Hubbard 1992). Their critique focuses on the influence of gender stereotypes on evolutionary explanations and is inspired by political and social values and ideals, such as gender equality, empowering women's life, and fighting patriarchal oppression. In this section, I report some of these criticisms targeting Sexual Strategies Theory and the theories it builds on. These criticisms show that feminist values gain cognitive relevance in at least four different ways.

The first criticism concerns Sexual Selection Theory. The primatologist and anthropologist Sarah Hrdy has argued that evolutionary theorists have ignored many widespread behaviours, shared by the majority of primate species.⁴⁵ For instance, scientists have paid little attention to data concerning competition among females, and females' engagement in promiscuous mating that is not necessary for fertilization. One of the reasons why this evidence was neglected was that it clashed with the post-Victorian prejudices and stereotypes about sex behaviours on which the picture of sexuality provided by Sexual Selection Theory was based. In this view, males are ardent and promiscuous lovers, and females are coy and monandrous (Hrdy 1986, 146). On the basis of these assumptions, evolutionists assumed that being sexually competitive was not advantageous for females and ignored all the behaviours related to this trait (Hrdy 1981/1999, 11).

However, Hrdy has argued that primatology provides extensive evidence showing that the assumptions that reproductive variance is greater among males than among females and that fertilization is the only reason females have to mate are unjustified (Hrdy 1986, 131). The prejudices and stereotypes underlying the traditional account have obstructed the attainment of an exhaustive understanding of sexuality, constraining research to analyze

⁴⁵ The study of nonhuman primates constitutes an important source of evidence for evolutionary psychology and biology. Since human beings and some nonhuman primates share a recent common ancestor in their evolutionary lineage, evolutionary scientists think that the study of these animals can shed light on the adaptive problems of our ancestors. For a critique of comparative analysis as source of evidence see Buller 2005 and section 2.1 of Chapter 1 of this dissertation.

females only in the role of mothers and as passive resources available to males (Fedigan 2001, 49). By neglecting data on the behaviour of female primates, Hrdy argues, the traditional account of sexuality gives only a partial explanation of sexual behaviour.

On a similar basis, Andrew Smiler (2011) has provided an extensive critique of Sexual Strategies Theory. As explained in the previous section, the theory predicts specific sex differences in the preference for mating strategies. In particular, it predicts that men have evolved a preference for short-term strategies and women a preference for long-term strategies. Women and men would then differ in the number of partners they desire and also in the features they find attractive in partners.

Evolutionary Psychologists dedicate an extensive analysis to the differences in the number of partners that men and women desire. For instance, Schmitt and colleagues emphasize how their studies show that men desire more short-term partners than women do (Schmitt et al. 2003). Moreover, according to Evolutionary Psychologists, their studies reveal that men and women significantly differ in the features they see as most important in potential mates (i.e. men would ascribe importance to attractiveness and women to resources, Buss & Schmitt 1993).

However, Smiler argues that the data collected in these studies show similarities between men and women, and not differences. For instance, Smiler reports that only a minority of the men tested by Schmitt and colleagues (2003) expresses a preference for short-term mating. In general, both women and men express a preference for having just one partner (Smiler 2011, 605; see also Tate 2011). Moreover, Smiler questions the claim that these studies would clearly show that women and men differ in the features they find desirable in partners. The studies conducted by Evolutionary Psychologists show that women and men agree on what are the most important features in partners, namely “dependable character, emotional stability/maturity, pleasing disposition, and mutual attraction/love” (Smiler 2011, 606). The characteristics on which Evolutionary Psychologists focus, namely attractiveness and status, are “ranked as being of middling importance (or lower)” in both women and men (Smiler 2011, 607).

It is worth noticing that the data showing these similarities are collected by the same proponents of Sexual Strategies Theory. Why do they focus on sex differences and neglect the substantial similarities revealed by their studies? According to Smiler, proponents of Sexual Strategies Theory “selectively focus on only those findings that support their theory” (2011, 604). Evolutionary Psychologists, then, would cherry pick data in order to corroborate their predictions and ignore the findings that might question the adequacy of their theory.

Let’s turn now to the feminist critique of Parental Investment Theory.

As previously mentioned, in the traditional view endorsed by Evolutionary Psychologists, males’ parental investment is rather small and, in principle, can amount to only one ejaculation necessary for conception. Female investment, instead, is huge and includes pregnancy, breastfeeding, and protection. This view has been widely criticized by feminist

researchers. For instance, Ruth Hubbard has challenged the traditional criteria to measure parental investment, in particular the role of anisogamy and the way gamete size is considered in this calculation. Scientists traditionally assume that size and quantity of gametes (i.e., small, “cheap” male gametes vs. big, “expensive” female gametes) are important for the evolution of human sexual behavior. However, Hubbard argues that although it is true that females “produce fewer eggs than males do sperm, and females gestate the embryos, [...] it is not obvious how to translate these facts into energy expenditures”, i.e., it is not clear in which degree these factors determine the amount of parental investment experienced by men and women (Hubbard 1992, 142). Moreover, Hubbard argues that considering size and quantity of gametes is not sufficient for an accurate calculation of parental investment and that an exhaustive evaluation of male parental investment should involve several aspects. Feminist philosopher of science Carla Fehr explains this point as follows:

“Eggs are larger than sperm. So, if one simply considers gamete size, male investment is smaller. However, males do not use a microscopic eyedropper to dispense one sperm at a time. Hubbard questions whether investment ought to be measured at the level of the individual gamete. When one includes the total amount of energy and resources that are need to produce sperm and spemen, the energy required to develop and maintain secondary sexual characteristics (differences between the sex that are not directly linked to the reproductive system), the costs of male-male competition, the costs that males of many species invest in defending territory, and the effort that males of some species put into parental care, male costs may turn out to be higher that researchers have historically expected.” (Fehr 2018)

These claims are based on research on primates’ parental investment. Feminist primatologists have shown that, contrary to what is traditionally argued by evolutionists, male parental investment is far from being insignificant for the survival of infants (Hrdy 1986, 143). In many species, fathers are extensively involved in direct care and protection of the offspring (Hrdy 1976; Hrdy 1986).

The feminist critique of Parental Investment Theory have important consequences for Sexual Strategies Theory too. Indeed, if feminist researchers are right in claiming that the traditional determination of the amount of parental investment is faulty, Evolutionary Psychologists should revise their account of sexual behaviour in humans (see also Vandermassen 2004; Vandermassen 2005; Ah-King 2007). In particular, if it is true that males' parental investment turns out to be greater than previously believed, Evolutionary Psychologists should revise their predictions about men's preference for short-term strategies.

Feminist researchers have also conducted various studies to test the predictions of Sexual Strategies Theory. For instance, Perrin et al. (2011) have designed three studies to test the assumption concerning the different desires of men and women in the context of a heterosexual romantic relationship. Findings do not support the differences predicted by Sexual Strategies Theory. By contrast, these studies highlight several similarities in the self-

reports of men and women concerning the desired loving behavior. They have detected just one clear gender difference: women's greater desire for mutual support and care in a relationship. Moreover, Pedersen et al. (2011) have tested the hypothesis that men prefer short-term strategies more often than women do and failed to confirm it, obtaining results that are inconsistent with Sexual Strategies Theory. Indeed, their studies show that preferences for long-term strategies are widely shared among men and women.

Feminist researchers have formulated compelling arguments against the traditional evolutionary explanation of human mating. Their critique has been instrumental in stimulating research that has helped to overcome various forms of gender prejudices and stereotypes, and to obtain a more inclusive and accurate picture of human sexual behavior.

Specifically, research inspired by feminist values has promoted the attainment of knowledge in four different ways. It has *raised sensitivity to evidence* that was disregarded because of gender bias (e.g. evidence on female primates' 'promiscuity' and male primates' parental investment) but that has primary importance in order to attain a clear understanding of human sexuality. Moreover, it has *challenged specific biases* in science by questioning the assumptions over the sexual competitiveness of women. Feminist research has also *increased the external consistency* of the evolutionary account of sexuality (i.e. coherence with related parts of science, such as studies in primatology). Finally, it has *provided an extensive evidence base* to test the predictive accuracy of this account (e.g. by conducting studies testing Sexual Strategies Theory's predictions).

Feminist values have a non-cognitive core, that is, the commitment to the emancipation of women, gender equality, freedom of patriarchal oppression, and empowerment of women's lives. However, these values are entangled with a commitment to fighting gender bias in science (e.g. neglect of relevant data, prejudices and stereotypes about the roles of sexes in sexual selection). They improved scientific research on human sexuality by contributing to the core activities of scientific reasoning - gathering and analyzing scientific evidence and assessing theories on the basis of that evidence - in an epistemically beneficial fashion.

The case study shows that feminist values helped with the attainment of knowledge. That is, they have played the cognitive role that is supposedly reserved for features of theories like internal consistency and predictive accuracy. The traditional distinction between cognitive and non-cognitive values, in which values are seen as either promoting the attainment of knowledge or promoting some ethical, political, social ideals, may then be too simplistic to account for this dual role that feminist values play. A dual characterization of feminist values in science seems more appropriate: non-cognitive values can have cognitive (promoting the attainment of knowledge) and non-cognitive (promoting ethical, political or social ideals) roles at the same time.

In the next section, I will discuss some similarities between cognitive and non-cognitive values and how these similarities can lead us to rethink the traditional

cognitive/non-cognitive distinction of values. I will also discuss how different non-cognitive values can have different effects on theory appraisal.

3. Similarities between cognitive and non-cognitive values

The case study presented in the previous sections sheds light on how feminist values have enhanced scientific practice and the assessment of evolutionary theories. It seems then that cognitive and non-cognitive values may play similar roles in theory appraisal, contrary to what is argued in the traditional view (see section 1 of chapter 1; see also Rooney 1992 and Rooney 2017).

One might object that, while the epistemic importance of cognitive values is robust and consistent, i.e., cognitive values are able to help with the attainment of knowledge in any context, the epistemically beneficial influence of non-cognitive values is ‘fortuitus’. For instance, that feminist values were able to improve the assessment of theories of the evolutionary account of sexuality is due to the widespread presence of sexist, androcentric assumptions in science (for an overview of the extensive influence of sexist and androcentric assumptions on science see for instance Schiebinger 2001 and Reed 1978). Were these assumptions absent, feminist values would be irrelevant to science. The distinction between cognitive and non-cognitive values would then remain somehow intact.

This argument does not constitute a successful strategy to defend the cognitive/non-cognitive distinction of values. While it is true that the role of non-cognitive values in promoting an objective and accurate assessment of theories might be dependent on factors like the presence or not of some assumptions in the scientific community, the same applies to cognitive values. As I have shown in Chapter 3, whether cognitive values are relevant to the assessment of theories and whether they can provide reasons to believe a theory true or reliable depend on certain contextual factors, such as the availability of evidence and the methodological assumptions made by the scientific community.

It should be noted that the relevance of non-cognitive values like feminist values may also depend on the field we consider. In some scientific disciplines, such as high-energy physics or pure mathematics, feminist values may be epistemically inert. In other words, in these fields, feminist values might be epistemically irrelevant to the appraisal of scientific theories. Nonetheless, this does not prove that non-cognitive values have no relevant role in these fields. Perhaps feminist values do not have any role, but other kinds of non-cognitive values may be relevant. For instance, Kent Staley has recently provided a case study showing that economic values played an important role in physics, in particular in the research concerning the Higgs boson (Staley 2017).

The same might be true for cognitive values. That is, the relevance of cognitive values for theory appraisal may be dependent on the field considered. For instance, while

simplicity as ontological parsimony may be epistemically relevant in some fields, such as physics (see for instance the discussion developed in Nolan 1997 and Baker 2003), simplicity may be irrelevant or even noxious in other fields, such as biology. An example is the epistemically detrimental effect of simplicity on the biological research aimed at categorize sexes, where the preference for ontologically parsimonious theories led to to miss the complexity involved in the matter (see for instance Sjöberg & Gentry 2007, 5 and Fausto-Sterling 1993).

Besides being epistemically relevant to theory appraisal and being contextually-dependent in their importance, there is a further similarity between cognitive and non-cognitive values that should be considered. As I explained in the introduction of this dissertation, cognitive values may differ from each other in the reasons of their desirability. While some are desirable because truth indicative, namely informative over the truth or empirical adequacy of theories, others are desirable because of pragmatic reasons. In the third chapter, I have argued that the reasons for the desirability of a value may depend on contextual factors. However, there are values that are more easily truth indicative than others. For instance, predictive accuracy bears a close relation to the truth or empirical accuracy of a theory, i.e., it is heavily loaded on the cognitive side (see section 1.1 of Chapter 1). By contrast, simplicity is a value which is much harder to classify. McMullin (1983) includes it in his list of epistemic values, but he notes a couple of problems, too. First of all, it is vague and at least partially subjective. Second, even when simplicity can be measured in a relatively exact way, e.g. in statistical model selection as the number of adjustable parameters, it is debatable whether it is a good guide to empirical adequacy. Foster and Sober made strong claims to this effect (Foster & Sober 1994), claiming that simpler models have a higher tendency to be empirically adequate. However, they had to weaken and to qualify these claims in later publications and to acknowledge the context-sensitivity of inferences where simplicity has epistemic value (see also Burnham & Anderson 2002). More precisely, a *ceteris paribus* preference for simpler models can help to reduce estimation error in statistical inference, but this does not justify the claim that simplicity is, in general, a reliable guide to truth or predictive accuracy.

The same is true for non-cognitive values. Different kinds of non-cognitive values may be more or less able to promote the attainment of knowledge. Let's consider a couple of examples.

One might argue that values as creationist or racist values may have the same epistemically beneficial influence that feminist values had in the case study analysed in this chapter. To the extent that these values promote a critical attitude towards established scientific theories, detect questionable assumptions, and lead to greater diversity in scientific discourse (e.g., Fuller 2007), their influence seems as legitimate as the influence of feminist values. Creation scientists have argued, for example, that confrontation with theories of divine creation and intelligent design will help biologists to develop a critical attitude toward

the otherwise uncontested foundations of the evolutionary paradigm, and thus put them into a position to discover potential problems with their preferred theory (see also the discussion in Laudan 1982). However, these values cannot play the cognitive role that feminist values have played in Evolutionary Psychology.

To understand this difference, one must appreciate that creationist, racist, and feminist values are not on a par with respect to their epistemic commitments (see also the discussion in Kourany 2010). For example, creationist values involve dogmatic views about the origin of the universe that rule out several astronomical and biological theories from the start. Racist values make similar commitments about essential differences between human races.

Feminist values are not committed to any dogmatic theory concerning (the absence of) differences between men and women, or the biological and psychological explanations of human mating. They do not prevent the pursuit of any biological or psychological research program. Rather, they propose a perspective for discussing and criticizing such research programs.

This is a substantial difference to creationist or racist values. Instead of accepting the geological evidence against a ‘young Earth’ and incorporating it into a more balanced picture, creation scientists developed immunization strategies against such data that allowed them to uphold the central values to which the research program is committed (namely the creation of life by divine intervention). The same can be said about science guided by racist values, such as the Aryan Physics (*Deutsche Physik*) of the 1920s and 1930s. That research program criticized the increasing mathematization of theoretical physics and called for a stronger emphasis on experimental research. However, it was severely compromised by the a priori commitment to the racial determination of natural science,⁴⁶ leading its proponents to reject ‘formalistic’ theories developed by scientists of Jewish descent (e.g., Einstein’s theory of relativity), in spite of the strong evidence in their favor.

Thus, while there is no intrinsic tension between feminist values and scientific method, creationist and racist values constrain, by the very nature of their commitments and the way they respond to evidence, the development and improvement of scientific research. This explains why feminist values can be described as having a cognitive role in theory appraisal, whereas such a description would be mistaken for creationist or racist values. A role for feminist values in science is then more legitimate than a role for racist values because they (intrinsically) have an epistemic dimension that aims at fighting bias, and not at trumping scientific evidence.

⁴⁶ See also the following quote from the preface of Lenard’s (1936, p. IX) textbook *Deutsche Physik*: “In reality science is, like all products of man, conditioned by race and blood. No people has ever endeavored natural science [Naturforschung] without standing on the fertile ground provided by pre-existing properties [sic] of Aryans” (translation by Jan Sprenger).

It should be noted that not all feminist values have this positive role. Indeed, ‘feminism’ is an umbrella term that denotes different and conflicting views, such as equity feminism and gender feminism (Sommers 1994). As Barry Kuhle (2012) argued, while equity feminism is not committed to any dogmatic view on men and women, gender feminism is. That is, gender feminism is committed to a theory that rules out any evolutionary explanations of psychological differences between men and women. By constraining the development of evolutionary explanations, gender feminism's influence on Evolutionary Psychology is similarly noxious as the influence of racist or creationist values. Equity feminism, on the other hand, does not involve such a dogmatic view. It asks for a fair treatment of women in Evolutionary Psychology, i.e., it asks for a careful study of women and men's psychological traits. Since equity feminism fights sexism, it is involved in a process of political change. Nonetheless, the aim of equity feminism is not formulating a ‘politically correct’ account of human mating. Rather, it aims to do justice to women by fighting bias in Evolutionary Psychology, and striving for empirically adequate explanations of human sexuality. Hence, equity feminism not only does not contrast with Evolutionary Psychology, but it also improves the reliability of its theories.

The similarities discussed in this section question the traditional distinction between cognitive and non-cognitive values. Not all the values included in the traditional lists of cognitive values promote the attainment of truth in any context, and not all the values traditionally seen as non-cognitive constitute an obstacle to this goal. Values can then be labelled as cognitive or non-cognitive depending on the *function* they perform in a specific context (e.g. Evolutionary Psychology), namely either promoting the attainment of knowledge or not. The distinction should then be reconceived by taking into account the actual function performed by values and the similarities between values discussed in this section.

4. Conclusion

In this chapter, I have made use of a case study to demonstrate the cognitive role that non-cognitive values can have in theory appraisal. I have argued that feminist values have enhanced the attainment of knowledge in Evolutionary Psychology by raising sensitivity to relevant evidence, challenging biased assumptions, improving the external consistency of theories, and providing evidence to test some of the central predictions formulated by Psychologists.

Moreover, I have argued that, since non-cognitive values can play the role that is supposed to be reserved to cognitive values, the classical cognitive/non-cognitive distinction needs to be reconceived. I have also pointed out that not any non-cognitive values can play the cognitive role that feminist values play in the case study analysed in this chapter. While

feminist values (at least the ones involved in equity feminism) are able to promote the attainment of knowledge, others (e.g. creationist, racist, gender feminism values) obstruct it or may do it in a limited fashion.

If feminists are right and the evolutionary account of human sexuality is fundamentally mistaken, one might ask whether the acceptance of these theories have any consequence for the society. If it is true that Evolutionary Psychology provides a sexist, biased account of human mating, does this have any consequence for lay people's understanding of sexuality?

The next chapter will explore how people judge possible mistaken decisions in science, such as the acceptance of false or inadequate theories. Philosophers of science argue that, when there is the chance of taking mistaken decisions in science (i.e. in cases of inductive risk), non-cognitive values help scientists weighing the possible non-epistemic consequences of errors, namely consequences with a moral, social, or economic dimension (see section 1.4 in Chapter 1). The next chapter focuses on understanding how lay people reason about inductive risk and possible consequences of errors, and whether and how non-cognitive values influence this reasoning.

Chapter 5

The psychology of inductive risk

An experimental study on the role of non-cognitive values in the assessment of cases of inductive risk⁴⁷

1. Introduction

Patients have been advised to take aspirin to prevent cardiovascular diseases for a long time (see ISIS-2 1988 and Steering Committee of the Physicians' Health Study Research Group 1989). Recent research has called into question the prescription of aspirin as primary prevention⁴⁸ to certain groups of the general population (e.g. Ridker et al. 2005). While cardiovascular diseases are among the primary causes of death among women (in particular in postmenopausal women), the majority of the studies tested the aspirin treatment on samples including only men (Lippman 2006, 11). Ridker and colleagues (2005) report that, of the five randomized trials testing aspirin as a primary prevention treatment for

⁴⁷ This chapter is based on joint work with Matteo Colombo and Leandra Bucher.

⁴⁸ While primary prevention treatments are interventions aimed at preventing diseases before they emerge (e.g. immunization against infectious diseases and safe lifestyle habits), secondary prevention measures aim at mitigating and controlling the effects of an ongoing disease (e.g. medical tests to detect diseases at early stages).

cardiovascular diseases, three tested only men, and of the 2042 vascular events studied, fewer than 180 occurred in women (Ridker et al. 2005, 1294; see also Hayden et. al. 2002). There are sex differences in the prevalence, severity, and symptoms of cardiovascular and several other diseases. There are sex differences also in the responses to medical treatments, which make collecting direct evidence of women's response to the treatment urgent. Including both women and men in clinical trials enables clinicians to better understand the safety and effectiveness of the treatment for the general population.

Deciding whether to accept or reject a scientific hypothesis, or to adopt or not a certain methodology, is always done in the face of uncertainty, which may result from mixed results, lack of relevant evidence, and scientists' disagreement over the reliability of their methodologies. Because of uncertainty, scientists may make mistaken decisions, such as accepting a hypothesis that should be rejected. For instance, on the basis of clinical trials including only males, scientists may accept the hypothesis that aspirin is a safe and effective treatment for cardiovascular diseases in the general population - though it may in fact be unsafe or ineffective for female patients.

Philosophers of science call the chance of taking wrong decisions in the face of uncertainty an *inductive risk*. Scientists may make errors in various stages of scientific research, such as when assessing whether a scientific hypothesis should be accepted, choosing to adopt a certain level of statistical significance or deciding to favour a certain interpretation of data over others (Douglas 2000; see also section 1.4 of Chapter 1). Some philosophers of science argue that when these possible mistakes have social, ethical or economical consequences, such as harming people's health by prescribing an unsafe medicine, non-cognitive values play an essential role in scientific reasoning.

Several examples from scientific practice and the history of science have been presented to show that some non-cognitive values are necessarily involved in the evaluation of inductive risks (e.g. Douglas 2000; Staley 2017). Less attention has been paid to the psychology of inductive risk, that is, to the reasoning and judgements of people confronted with possible inductive risks. This lack of attention is surprising. After all, non-cognitive values are features of human psychology, and clarifying how they relate to people's understanding of scientific reasoning and science-based policies would be congenial to recent calls from scientists, philosophers, and policy-makers to better align scientific research and innovation with the values, needs, and expectations of society (e.g. Figueiredo Nascimento et al. 2016; Schroeder 2017).

After having engaged with case studies and philosophical arguments about the role of values in science, in this chapter I turn to describe and discuss the results of an empirical study I have conducted on the psychology of inductive risks.

In the first section, I present the argument from inductive risk. I specify the types of error that scientists may incur in cases of inductive risk, as well as the roles that non-cognitive values play in the evaluation of the consequences of errors. In the second and third sections,

I present my empirical study and its results. In this study, my co-authors and I have investigated how the type of error, the scientific domain in which an error may be made, and the consequences of errors, predict lay people's judgements about possible cases of inductive risk. Moreover, we also probed the association between personal features of our experimental participants (i.e., their sex, race, and political orientation) and their evaluation of possible cases of inductive risk. In the fourth section, I discuss the results and the contributions of the study to the philosophical and psychological literature. In the fifth section, I summarize the study.

2. Inductive Risk: The argument and psychological evidence

When assessing a scientific hypothesis, scientists may end up in one of these scenarios:

- (1) The hypothesis is accepted and it is in fact true.
- (2) The hypothesis is rejected and it is in fact false.
- (3) The hypothesis is accepted but it is in fact false.
- (4) The hypothesis is rejected but it is in fact true.

While in the first two scenarios no error is involved, the third and the fourth involve two types of error. The third scenario involves a type I error or false positive, i.e., a hypothesis that should be rejected, because actually false, is accepted as true. The fourth scenario instead involves a type II error or false negative, i.e., a hypothesis that should be accepted, because actually true, is rejected as false. Scientists incur these errors because the assessment of hypotheses is always done in the face of uncertainty (Hempel 1965, 83).

Errors are undesirable for various reasons. They are undesirable because they constitute epistemic failures. When accepting an actually false hypothesis as true scientists fail to attain knowledge. Moreover, errors are undesirable because they may have *non-epistemic consequences* (Douglas 2000, 559), namely consequences with a social, moral or economic dimension. For instance, accepting the hypothesis that a medical treatment is safe when it is actually not safe may have the non-epistemic consequence of doing a moral wrong, such as harming or killing people. Because of the possibility of non-epistemic consequences, some philosophers have argued, non-cognitive values play a fundamental role in cases of inductive risk. Specifically, non-cognitive values would provide scientists with social, moral, and economic standards for comparing and weighing the possible consequences of errors.

The argument from inductive risk challenges the value-free ideal of science by showing that non-cognitive values play a necessary role in scientific reasoning and that such a role is often legitimate and beneficial (Rudner 1953; Hempel 1965; Douglas 2000). According to Richard Rudner, in cases of uncertainty, non-cognitive values should determine whether the available evidence is strong enough to warrant the acceptance (or rejection) of a

hypothesis, namely whether the available evidence provides scientists with a relatively high degree of confidence on the truth/falsity of the hypothesis (Rudner 1953, 3). The idea is that, to draw a conclusion about whether a given hypothesis should be accepted as true or not, evidence should be strong enough to ensure that scientists avoid errors. The possible moral, economic, and social consequences attached to those errors help determine when the evidence is “strong enough”.

Committing an error may be more or less *serious* depending on how bad its consequences are. Different levels of evidential strength may be required, depending on how undesirable the possible consequences are. To illustrate this point, consider the examples provided by Rudner:

“if the hypothesis under consideration were to the effect that a toxic ingredient of a drug was not present in lethal quantity, we would require a relatively high degree of confirmation or confidence before accepting the hypothesis – for the consequences of making a mistake here are exceedingly grave by our moral standards.” (Rudner 1953, 2)

Accepting a hypothesis as “the toxic ingredient is not present in lethal quantity” requires strong evidence pointing in that direction. We want to be sure, Rudner argues, that the evidence convincingly shows that the hypothesis is true. This is because a possible bad consequence of a false positive (accepting the hypothesis that the toxic ingredient is not present when it is actually present) is harming or killing people. This consequence is particularly undesirable for our moral standards (provided that we endorse a certain moral view for which killing is particularly wrong) and we want to be sure that by accepting the hypothesis we will not face this consequence.

Compare this case with the second example provided by Rudner, which is a case from industrial quality control:

“On the other hand, if say, our hypothesis stated that, on the basis of a sample, a certain lot of machine stamped belt buckles was not defective, the degree of confidence we should require would be relatively not so high.” (Rudner 1953, 2)

Making a mistake in this case has less serious consequences than in the medical case, since accepting the false hypothesis that the lot is not defective has relatively little chance of killing people. Accepting the false hypothesis that the lot is not defective might have the consequence – among others – of producing malfunctioning belt buckles, and frustrating the desire of customers of buying working belt buckles. However, this consequence is less harmful than the consequence of killing people with a toxic ingredient. Hence, because of the different levels of undesirability of the possible consequences, the strength of the evidence to warrant the acceptance of a hypothesis will be less demanding in this case than in the medical case. The degree of confidence we need in order to accept (or reject) a scientific hypothesis is then dependent on the risk of being wrong we are willing to take

(Rudner 1953, 3).

Besides determining how strong the evidence should be to warrant a conclusion about scientific hypotheses, non-cognitive values determine which of the possible errors is more tolerable. Douglas (2000) explores this point when discussing the possible errors involved in the choice of the level of statistical significance in toxicological studies.⁴⁹ Specifically, she focuses on research studying how dioxin can induce cancer. These studies aim at determining whether the laboratory animals exposed to doses of dioxin “exhibit significantly more cancer than the control animals” (2000, 567), namely the animals that are not exposed to the doses:

“The statistical comparison between exposed and control animals is particularly important for cancer studies, where both control animals and exposed animals will usually exhibit some cancer. What must be determined is whether the exposed animals exhibit significantly more cancer than the control animals. Only when a cancer rate is significantly different from the control cancer rate is it considered a genuine result of the dosing. Thus, setting the standard for statistical significance will impact what is considered a response caused by the dosing.” (Douglas 2000, 566-7)

Deciding the level of statistical significance means determining in which cases there is a large difference in cancer rates between exposed animals and control animals. The tests may incur false positives and false negatives, and scientists, when deciding the level of statistical significance, have to consider which kind of error they are more willing to tolerate (Douglas 2000, 566). More precisely, scientists have to find a balance between false positives and false negatives that fits their desire to avoid a certain kind of error.⁵⁰

If scientists prefer to avoid false positives, they may decide to apply a stringent standard for statistical significance. In this case, the response is considered significant when there is a greater difference in cancer rates between exposed animals and control animals. This means lowering false positives and increasing false negatives.

If scientists, instead, judge false positives as less undesirable than false negatives, they may decide to apply a less strict standard for statistical significance. Applying a less strict standard increases the likelihood of false positives and lowers the false negatives (Douglas 2000, 567).

Now, why would scientists want to avoid a certain kind of error? Whether an error is less undesirable than others depends on its social, moral or economic consequences, and these consequences are evaluated and compared on the basis of non-cognitive values.

⁴⁹ Douglas' example does not concern the assessment of scientific hypotheses, but the possibility of inductive risk in the choice of methodologies. As discussed in the introduction, the assessment of scientific hypotheses is just one of the stages in which scientists face inductive risk (Douglas 2000, 565).

⁵⁰ Douglas specifies that it is not possible to lower both types of error. To do so, a more precise test should be developed (Douglas 2000, 566).

Consider Douglas' example:

“false positives are likely to lead to stronger regulation than is warranted (or overregulation); false negatives are likely to lead to weaker regulation than is warranted (or underregulation). Overregulation presents excess costs to the industries that would bear the costs of regulations. Underregulation presents costs to public health and to other areas affected by damage to public health.” (Douglas 2000, 567)

Depending on how the possible consequences are perceived (e.g. morally bad or not), scientists may prefer a certain level of statistical significance. If, for instance, scientists think there is the risk of damaging public health and they see this as a particularly undesirable consequence (and less tolerable than having onerous costs of regulations), they will opt for a less strict level of statistical significance. Depending then on the non-cognitive values endorsed and what is valued the most by the society or scientific community, either public health or low costs of regulations,⁵¹ scientists may take different decisions over statistical levels of significance.

Non-cognitive values, then, have a role in determining which consequences should be avoided and which error is more tolerable. This does not mean that non-cognitive values *determine* whether a hypothesis is true or whether a level of statistical significance is more epistemically justified than others. What the argument from inductive risk shows is that scientific and non-epistemic factors are intertwined in science: decisions over scientific hypotheses, methodologies, and levels of statistical significance may involve ethically, economically, and socially relevant issues and non-cognitive values are necessary to handle these issues.

Philosophers of science have drawn on several historical case studies to clarify the notion of inductive risk and the importance of non-cognitive values in scientific practice (e.g. Douglas 2000; Staley 2017). However, relatively little is known on how non-cognitive values might influence people's reasoning in cases of inductive risk. The topic of inductive risk has not received much attention in psychology either, though psychologists have conducted several studies on how people's prior beliefs and values affect their reasoning and judgments under uncertainty.

First of all, existing research in social psychology provides us with results that are relevant to understanding *how people react to uncertainty in science*. For instance, Lord, Ross and Lepper (1979) have studied how people react when presented with mixed or inconclusive findings. Their study revealed that uncertainty in results does not make people reconsider their own beliefs. Rather, uncertainty in evidence induces people to make their beliefs even stronger, namely to *belief polarization* (Lord, Ross & Lepper 1979, 2103).

⁵¹ This is a simplified scenario. The differences in the consequences of possible errors may be more or less marked. For instance, a type I error may have social or moral (e.g. harming people's health) as well as economic consequences (e.g. losing money).

Second, existing results in cognitive psychology are relevant to understanding *how personal values influence the way evidence, scientific methodologies, and practices are evaluated*. Koehler (1993) found that the way scientists assess the *quality of evidence* is affected by their own values and prior beliefs. Specifically, scientists would tend to evaluate the studies supporting their prior beliefs as “more relevant, methodologically sound, and clearly presented than the identical studies that opposed their beliefs” (Koehler 1993, 46; see also Munro & Ditto 1997).

Research in social psychology has also revealed that values influence how people assess the *risks and benefits of practices*. Kahan and colleagues (2010), for example, found that, depending on the political values endorsed, people emphasize either the risks or the benefits of scientific practices. Their study has focused on understanding how cultural and political values affect people's assessment of the risks and benefits of HPV vaccination. While people endorsing individualist-hierarchist values see the vaccination as encouraging unprotected sex and ineffective in preventing cancer, participants endorsing communitarian-egalitarian values⁵² understand evidence on the vaccine as showing its effectiveness and positive impact on society.⁵³

Taken together, these studies and other psychological studies reveal that personal values play a distinctive role in both scientists' and lay people's reasoning in cases of uncertainty. Their results independently corroborate philosophers' criticism of the value-free ideal on the descriptive level. Science is not a value-free enterprise: values do influence how scientists reason about evidence and scientific methodologies. The study I conducted and I am about to present contributes to this line of research, by explicitly bringing the psychology and philosophy of inductive risk together for the first time.

3. Experimental study on inductive risk

Philosophers of science like Rudner and Douglas have argued that the evidential strength necessary to warrant the acceptance of a hypothesis may vary across contexts, depending on

⁵² Kahan and colleagues explain these positions as follows. Individualist-hierarchists support a society in which individuals “are expected to secure the conditions of their own well-being without interference or assistance from the collective” and a social order in which “all manner of prerogatives and obligations are assigned in a stratified way on the basis of relatively stable individual characteristics, such as gender, ethnicity, and class” (Kahan et al. 2010, 504). Communitarian-egalitarians value a society in which “the needs of the collective take precedence over those of the individual and in which the collective is deemed responsible for securing the conditions of individual flourishing” and a social order which factors as gender, ethnicity, and class are irrelevant to the distribution of status and entitlements (ibid.).

⁵³ Several studies have reported similar results. See for instance Liu and Ditto 2013 on how people, depending on the moral values endorsed, emphasize either the risks or the benefits of capital punishment.

the seriousness of the possible non-epistemic consequences of errors. The first question explored in our study, then, is:

(Q1) Do lay people's judgments about possible inductive risks vary across different domains such as the domains of healthcare, education, and car design?

We predicted that people will make different judgments about possible inductive risks in different domains. Based on the philosophical arguments presented in the previous section (e.g. Rudner 1953), we expected that the non-epistemic consequences of a possible error in some domains would be perceived as more serious than the consequences of a possible error in other domains. In particular, we hypothesized that, when confronted with a policy decision based on a possible false positive, lay people will be more uncertain that the decision was a good one in the healthcare domain than in the education domain, and that lay people will be more uncertain that the decision was a good one in the education domain than in the car design domain.

Our second question is:

(Q2) Do lay people's judgments about potential inductive risks vary as a function of whether or not the possible risks concern oneself (as opposed to members of other racial or sexual groups)?

We predicted an affirmative answer to this question. This prediction was based on existing psychological results on motivated reasoning, confirmation bias, and risk perception. Studies in psychology show that people evaluate the reliability of scientific research in a biased way. For instance, some studies show that whether a study is assessed as convincing or not may depend on whether it has personal implications for the people assessing it. Kunda (1987) has shown this difference in reactions in a study in which caffeine addicted and non-caffeine addicted assessed a study reporting serious health risks in the use of caffeine. The study shows that heavy caffeine consumers were more uncertain about the reliability of the study than low caffeine consumers (Kunda 1987; see also Kunda 1990). Moreover, some risk perception studies show that the way individuals perceive risks not only depend on the features of the risk (e.g. new vs. old risk, number of people exposed), but also on the personal features of the perceivers, such as their beliefs, values, and worldviews. For instance, Slimak and Dietz (2006) conducted a survey in which they asked participants how concerned they were about various ecological risks (e.g. acid rain, ozone hole) and to rank them in order of importance. Their results show that people endorsing different values (e.g. altruism vs. self-interest) and with different levels of expertise in ecological matters ranked ecological risks differently (Slimak & Dietz 2006, 1695).

Third, our study clarifies asymmetries between judgements in the face of false

positives versus false negatives, by asking:

(Q3) Do lay people's judgments about potential inductive risks vary as a function of whether the potential error consist in a false positive or a false negative?

We predicted that false positives would generally be judged as more serious than false negatives. This prediction was based on suggestions in the statistical literature, which often presents incurring type I errors as a worse scenario than incurring type II errors in hypothesis testing (Nickerson 2000, 243; Wasserman 2013, 150). To give an example, when uncertain about the safety of a medicine, deciding to prescribe people with the medicine (when it is actually an unsafe medicine) would be worse than deciding to not to prescribe it (when it is actually a safe medicine). This is because the consequences of the former error (e.g. harming people with an unsafe medicine) are seen as more serious than the consequences of the latter (e.g. preventing people from having access to a safe medicine). However, the assumption that false positives are always more serious than false negatives has been criticized as simplistic (for an introduction on this criticism see for instance Smith 2013). The reason is that the consequences that errors have vary across contexts. For instance, incurring a false negative may be particularly undesirable when this involves denying access to a cure for a widespread disease with high fatality rate. To determine whether it is better to incur a type of error rather than the other, critics argue, involves considering the actual consequences that errors have in a specific context.

Rudner's and Douglas' arguments focus on scientists, and not on lay people, and scientists and lay people may substantially differ in the ways they react to inductive risks. Scientists have a privileged epistemic position, having a better understanding of the disputes about the reliability of scientific methodologies and generalizability of results. The results of our study, then, are not generalizable, since they may be unrepresentative of how experts react and perceive possible cases of inductive risks, and also of how non-experts in countries other than the USA (i.e., the country of residence of our participants) may perceive inductive risks across the three domains we considered. However, the decision to accept or reject a scientific hypothesis affects both scientists and lay people. Understanding how lay people react to uncertainty in science is important in order to meet the recent calls made by institutions like the European Union and the American Association for the Advancement of Science (AAAS), which have implemented various projects aimed at aligning scientific research and innovation with citizens' values, needs, and expectations (e.g. Figueiredo Nascimento et al. 2016).

Despite these limitations, our results contribute fresh understanding of how lay people judge possible cases of inductive risks, and how their judgements are associated with personal, non-cognitive values. The contribution of the present study is important for the work of philosophers of science, psychologists, and policy makers, as I will explain in

Section 4 below. In particular, it is relevant to the discussion of the attainability of the value-free ideal and to discuss how and whether scientific research should be aligned with citizens' values, needs, and expectations.

In the next sections, I will present the method used in our study and report the results. I will first present the results concerning people's judgments of possible type I errors in section 3.1 and the results concerning people's judgments of possible type II errors in section 3.2.

3.1 Experiment: Error Type I

The experiment was conducted online, utilizing the Qualtrics Survey Software within the Amazon Mechanical Turk (MTurk) environment. We only allowed MTurk participants with an approval rate > 95% and with a number of HITs approved > 5000 to submit responses. The instruction and the material was presented in English.

Participants

A total $N = 374$ ⁵⁴ participants took part in the Error Type 1 experiment, mean age 32.93 years, $SD = 9.078$. The participants received a small monetary payment. Each participant rated three vignettes, one for each of the three domains of research we used (i.e., healthcare, education, and car design). Approximately half of the participants ($n = 182$) were presented with and made judgements about studies including samples composed of only "minorities" (i.e., women, Black, and left-handed people). The other half of the participants ($n = 192$) was presented with and made judgements about studies including samples composed of only "no minorities" (i.e. men, White, and right-handed people). The combinations of research domain and sample were counterbalanced across participants to meet the criteria of a complete design.

Design, Material, and Procedure

The experiment follows a $3 \times 3 \times 2$ factorial design, with the factors Domain of Research (health, education, car design), Sample (gender, skin color, handedness), and Minority Group (no, yes) (see Figure 1).

Each vignette was composed of three sentences, where the content of the individual vignettes varied in accord with the factorial design of the experiment. The first sentence introduced a scenario where scientists had tested a certain novelty, within a certain domain

⁵⁴ Of the $N = 374$ participants, $n = 129$ were female, $n = 242$ male, and 3 diverse; $n = 220$ were "White", $n = 21$ "Black", $n = 106$ Asian, $n = 17$ Hispanic (or Latino), and $n = 10$ of other ethnicity, $n = 327$ of the participants stated that they were right-handed, $n = 40$ left-handed, and $n = 7$ described themselves as ambidextrous.

of life (e.g. a new medical drug in the research domain of "healthcare"). However, the sample used in the study that tested the novelty was not suitable to represent the general population. Instead, the sample was biased, focusing on a special "aspect" of the population (in terms of either "gender", "skin color", or "handedness"). Further, the sample included either only representatives of the "minority" or "no minority" group within the categories of gender, skin color, and handedness, i.e. women vs. men, Black vs. White, and left-handed vs. right-handed people).

Experimental Design					
	Domain of Research	health care	education	car design	Minority Group
Sample	gender	men	men	men	no
		women	women	women	yes
	skin color	white	white	white	no
		black	black	black	yes
	handedness	right	right	right	no
		left	left	left	yes

Figure 1. The $3 \times 3 \times 2$ factorial design of the experiment, with the factors Domain of Research (health, education, car design), Sample (gender, skin color, handedness), and Minority Group (no, yes).

[First sentence of the vignette]:

"A group of scientists recently tested [the safety of a new medicine⁵⁵ / the validity of a new method to measure intelligence⁵⁶ / the appeal of a new car⁵⁷] on a group including only [women⁵⁸ / men⁵⁹ / Black people⁶⁰ / White people⁶¹ / left handed people⁶² / right handed people⁶³]."

⁵⁵ Domain of Research: healthcare.

⁵⁶ Domain of Research: education.

⁵⁷ Domain of Research: car design.

⁵⁸ Gender sample with representatives of the "minority group".

⁵⁹ Gender sample with representatives of the "no minority group".

⁶⁰ Skin color sample with representatives of the "minority group".

⁶¹ Skin color sample with representatives of the "no minority group".

⁶² Handedness sample with representatives of the "minority group".

⁶³ Handedness sample with representatives of the "no minority group".

The second sentence of the vignette highlighted the uncertainty, raised by scientists, about the adequacy of the scientific test due to the unrepresentative sample, used for the test:

[Second sentence of the vignette]:

"The broader scientific community disagrees about the adequacy of this test, since they are uncertain about the generalizability of the test results to the general population".

The third sentence reports about the decision *in favor of a change* (bearing the risk for committing a Type I-error):

[Third sentence of the vignette]:

"Despite the uncertainty, the decision was made to introduce [the new medicine to the market⁶⁴ /the new method to measure intelligence into examinations for college admissions⁶⁵ / the new car to the market⁶⁶]."

Participants were asked to rate three vignettes (one from each domain of research), one by one. Subsequent after reading a vignette, participants were asked to consider "How certain are you that the decision was a good one?" The ratings of the participants were collected, using a 7-point scale with the extremes (1) "Very certain" and (7) "Very uncertain".

The role of personal features and non-cognitive values in the assessment of inductive risk

A further aim of our study was to examine whether and to what extent personal features and non-cognitive values play a role in the assessment of inductive risk. To that end, we embedded a number of personal features as "co-variables" into our analyses of the ratings, and assessed to what extent the sex of our participants (male, female, diverse), the race (White, Black or African American, Asian, American Indian and Alaska Native, Native Hawaiian and Other Pacific Islander, or some other race⁶⁷), and handedness (left, right, ambidexter) modulated the assessment of inductive risk in our various conditions.

Further, we assessed a number of non-cognitive values, namely "general political interest"⁶⁸, the "general political attitude"⁶⁹ in terms of "conservative", "moderate", or

⁶⁴ Phrasing in the "domain of healthcare" vignette.

⁶⁵ Phrasing in the "domain of education" vignette.

⁶⁶ Phrasing in the "domain of car design" vignette.

⁶⁷ Race categories according to the "United States Census Bureau",
<https://www.census.gov/mso/www/training/pdf/race-ethnicity-onepager.pdf>

⁶⁸ To assess "general political interest", participants answered the question "How interested would you say you are in politics?", using the following scale: 1 = very, 2 = quite, 3 = a little, 4 = hardly, and 5 = not at all.

⁶⁹ To assess "general political attitude, participants answered the question "Generally speaking, do you

“liberal”, and specific attitudes towards “race related social issues” and “gender related social issues”, each in terms of “conservative”, “moderate”, or “liberal”^{70,71}.

In additional analyses that included ratings of the general political interest, general political attitude, and the two specific political attitudes as covariates, we examined the potential influence of “non-cognitive values” of our participants on the assessment of inductive risk of error type I and II.

Why these vignettes

For our study, we decided to focus on inductive risk in the healthcare, education, and car design domains for several reasons. First, errors’s seriousness varies depending on which of the three domain is involved. Compared to the car domain, inductive risks in the medical and educational domains are more serious.

Take the medical vignette. It concerns the exclusion of some groups from the clinical trials testing the safety of a new medicine, and the disagreement among scientists over the generalizability of the results to the general population. In this scenario, a wrong decision may have the serious consequence of harming the health of some groups of the population. Compare this scenario to the car design domain. Here a mistaken decision may have the consequence - among others - of producing an aesthetically unappealing car and making little profit. This consequence is arguably less serious or undesirable than the consequence of harming or killing people with an unsafe medicine (provided that we endorse a certain moral view for which killing is worse than making little profit).

A second reason for focusing on these domains is that scientists do face cases of inductive risks in these areas. In our vignettes, the exclusion was based on either the sex or race of participants. Research in the biomedical and clinical spheres has been widely criticized for excluding certain groups of the population from clinical trials and for the consequences that such exclusion may have for the health of the groups excluded (see also Labots et al. 2018). For instance, women have been excluded from clinical trials for a long time (as in the case of aspirin mentioned in the introduction), and the decision of excluding them was taken on the basis of various ethical and economic concerns. Impeding their participation was seen as a measure to avoid possible tragic consequences for fetuses.⁷²

usually think of yourself as conservative, moderate, or liberal on a scale 1 = (very) conservative, 2 = moderate, 3 = (very) liberal.

⁷⁰ To assess the attitude towards “race related social issues”, participants answered the question “With regard to race related social issues, do you usually think of yourself as conservative, moderate, or liberal?”. Answers very indicated on a scale, 1 = (very) conservative, 2 = moderate, 3 = (very) liberal.

⁷¹ To assess the attitude towards “gender related social issues”, participants answered the question “With regard to race related social issues, do you usually think of yourself as conservative, moderate, or liberal?”. Answers very indicated on a scale, 1 = (very) conservative, 2 = moderate, 3 = (very) liberal.

⁷² The practice of excluding women from clinical trials was implemented after that the decision to prescribe women with thalidomide in the 1960s revealed to be a mistake (Lippman 2006; Liu & Dipietro

Moreover, women, because of hormone fluctuations, were seen as very expensive to test (Liu & Dipietro Mager 2016, 2).

Non-white people have also been (and still are) excluded from clinical trials. For instance, HPV vaccines, in particular Gardasil and Cervarix, have been tested mainly on White women (Intemann & de Melo-Martín 2010). However, as Kristen Intemann and Inmaculada de Melo-Martín report, there are significant differences in the frequency White women, Hispanic, and Black women are diagnosed with cervical cancer, in particular “hispanic women are diagnosed with cervical cancer almost twice as often, and African American women more than 1.5 times as often as non-Hispanic white women” (Intemann & de Melo-Martín 2010, 209). The ability of the studies to account for the differences in frequency of cervical cancer and produce vaccines that can effectively prevent cervical cancer in women of different ethnicities has then been questioned.

The decision of excluding certain groups from clinical trials was based on the assumption that a certain subject, namely a male/White subject, can serve as representative of the species (see Simon 2005, 1517). However, this assumption has been questioned, since there are significant physiological differences between the sexes and races in health risks and responses to medical treatments (Pinn 2003; Lippman 2006; Valles 2012). Testing medicine and treatments on only certain groups of the population has the undesirable consequence of depriving other groups (e.g. women, Black people) of safe and effective cures (Lippman 2006; Kukla 2016).

A further reason for testing people’s sensitivity to possible cases of inductive risks in these domains was that personal features, namely sex and race, and the political values endorsed may interact with the variables manipulated in the vignettes.

For instance, the education vignette concerned the adoption of a certain measure of intelligence to be used in tests for college admission. As in the medical scenario, the measure was tested by focusing on certain groups of the population and excluding others. In this scenario, mistaken decisions may have important non-epistemic consequences (e.g. limiting educational opportunities) and whether these consequences are seen as undesirable may depend on personal demographic features and values. Studies have reported for a long time differences in the performance of students of different ethnicities. In particular, a gap was reported in the performance of White and Black students (e.g. Coleman et al. 1966): “a simple comparison of mean test scores typically finds black students scoring roughly 1 standard deviation below white students on standardized tests” (Fryer & Levitt 2004, 447). Researchers have suggested various explanations to account for this gap, such as biological differences, the influence of racial bias on teachers’ perceptions of the performance of

Mager 2016). There was no evidence available on the possible risks and benefits of the treatment for women. However, thalidomide was prescribed to pregnant women to alleviate morning sickness. The consequences were tragic: thalidomide blocked the development of limbs in fetuses and almost 1000 children were born with malformations. Only half of them survived.

students, differences in school quality, and the effect of family structure and poverty on students' educational achievement (e.g. Card & Rothstein 2004; Fryer & Levitt 2004). We expected that participants of different ethnicities may differ in their reactions to the decisions of excluding groups (e.g. Black people) from the testing of measures of intelligence.

Results

Participants' ratings⁷³ were submitted to a univariate factorial analysis (UNIANOVA) with the factors Domain of Research (health, education, car design), Sample (gender, skin color, handedness), and Minority Group (no, yes).

The analysis revealed a main effect of Sample, $F(2, 17) = 6.869$, $p = .001$; $\eta^2_{\text{part}} = .012$, a two-way interaction of Domain of Research $\times$ Sample, $F(4, 17) = 3.072$; $p = .016$; $\eta^2_{\text{part}} = .011$, a two-way interaction of Sample $\times$ Minority Group, $F(2, 17) = 4.917$; $p = .007$; $\eta^2_{\text{part}} = .009$, and a three-way interaction of Domain of Research $\times$ Sample $\times$ Minority Group, $F(4, 17) = 4.125$; $p = .003$; $\eta^2_{\text{part}} = .015$.

Post-hoc tests (Bonferroni-corrected) were conducted and revealed different patterns of ratings in the different research domains.

Research domain of healthcare

In the domain of "healthcare", ratings signal significantly higher certainty (that the decision was a good one) with "handedness" samples ($M = 4.00$; $SD = 1.90$) as compared to "gender" samples ($M = 4.98$; $SD = 1.83$; $t(242) = 4.066$; $p < .001$; $d = -.528$) and "skin color" samples ($M = 4.68$; $SD = 1.92$; $t(241) = 2.810$; $p < .001$; $d = -.36$). Ratings of "gender" and "skin color" samples did not differ significantly ($p = .59$).

Ratings of vignettes presenting the samples of right-handed people ($M = 3.46$; $SD = 1.70$) expressed significantly higher certainty (that the decision was a good decision), compared to ratings of vignettes presenting the sample of left handed-people ($M = 4.55$; $SD = 1.94$; $F(1, 112)$; $p = .002$; $\eta^2_{\text{part}} = 0.85$).

Ratings of vignettes presenting "minority groups" (women, Black people, and left-handed people), did not differ significantly from each other ($p = .74$). However ratings of vignettes presenting the "no minority" groups (men, White people, and right-handed people) expressed different degrees of certainty (that the decision was a good one) as follows.

The ratings expressed the highest certainty with samples of "right handed people" ($M = 3.46$; $SD = 1.70$), and the least certainty with samples of "men" ($M = 5.12$; $SD = 1.74$). Participants were significantly more certain that the introduction of a new drug to the market was a good decision, when the drug had been tested with a sample of right-handed people,

⁷³ With the scores as follows: 1 = Very certain; 2 = certain; 3 = Slightly certain; 4 = Neither certain, nor uncertain; 5 = Slightly uncertain; 6 = Uncertain, and 7 = Very uncertain.

compared to when the drug had been tested with a sample of only White people ($M = 4.82$; $SD = 1.95$; $t(116) = 4.096$; $p = .001$; $d = 0.74$), or with a sample of only men ($t(133) = 5.268$; $p = .001$; $d = 0.96$). Figure 2 illustrates the results obtained for the “domain of healthcare” scenarios.

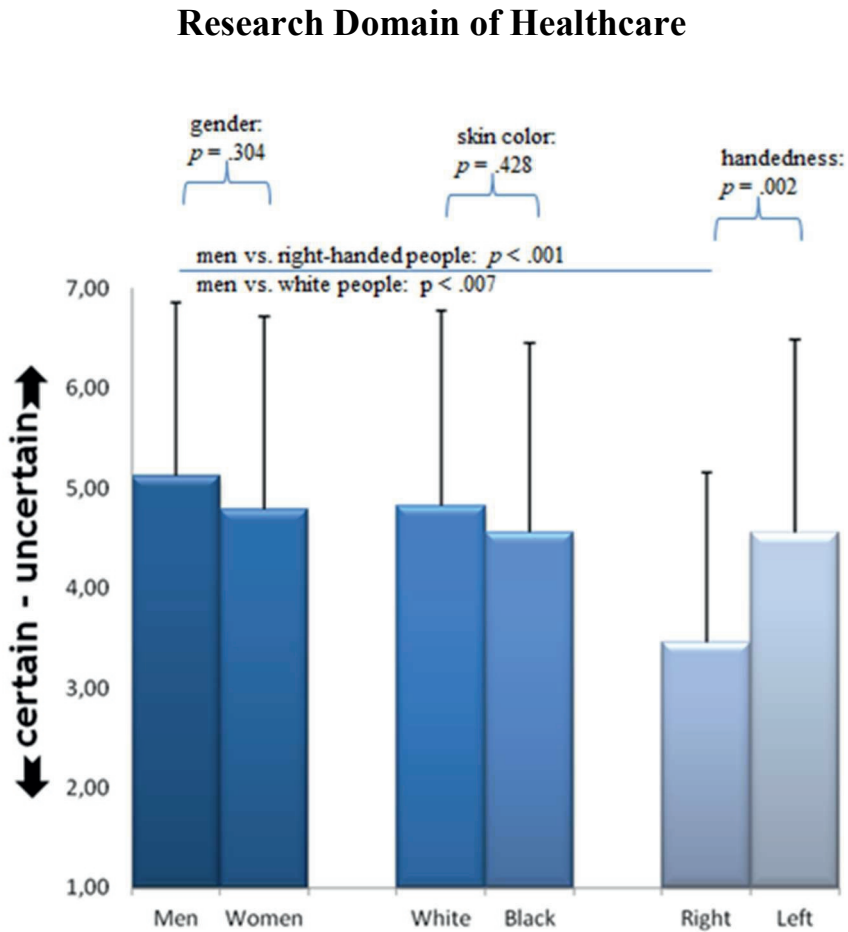


Figure 2. In the domain of healthcare, the ratings expressed significantly higher certainty (that the decision was a good one) with “right-handed people” samples compared to samples of “men”, “White people”, and “left-handed people”. Error bars show standard deviations.

Research domain of education

Ratings of “education domain” scenarios indicate that the participants were more certain to allow a new method to measure intelligence into examinations for college admissions when the new method had been tested on a sample of only women ($M = 4.61$; $SD = 1.84$), compared to when it was tested on a sample of only men; ($M = 5.15$; $SD = 1.70$; $F(1, 147) = 3.490$; $p = .064$; $\eta^2_{\text{part}} = .023$). Further, participants were more certain when the method had been tested on a sample of only right-handed people ($M = 4.40$; $SD = 1.926$); compared to only left-handed people ($M = 5.27$; $SD = 1.567$; $F(1, 112) = 6.835$; $p = .010$; $\eta^2_{\text{part}} = .058$). Ratings did not significantly differ for samples of only White people ($M = 4.73$; $SD = 1.91$) and samples of only Black people ($M = 4.21$; $SD = 1.73$) ($p = .131$).

Comparisons of ratings of the “minority group” samples (women, Black people, and left-handed people) showed that participants were more uncertain (that the decision was a good one) when decisions were based on tests conducted with samples of only left-handed people, compared to tests based on samples of only Black people ($t(111) = 3.261538$; $d = 0.642$). (All other $ps > .10$).

Comparisons between the “no minority group” samples showed that participants were marginally significantly more uncertain (that a decision was a good one) when decisions were based on samples of only men, compared to decisions based on samples of only right-handed people ($t(134) = 2.358$; $p = .059$; $d = 0.413$). (All other $ps > .55$). Figure 3 illustrates the results obtained for the “domain of education”.

Research domain of car design

Participants expressed marginally significantly higher uncertainty (that the decision to introduce a new car to the market was a good one) when the design of the car had been evaluated by samples of only women ($M = 5.0$; $SD = 1.445$), compared to when the car design had been evaluated by samples of only men, ($M = 4.65$; $SD = 2.044$), $F(1, 129) = 3.019$; $p = .085$; $\eta^2_{\text{part}} = .023$).

Comparisons of ratings of the “minority group” samples (women, Black people, and left-handed people) showed that participants were more uncertain (that a decision was a good one) when the decision was based on tests including samples of only women, compared to samples of only left-handed people ($M = 4.40$; $SD = 1.891$; $t(124) = 2.5316$; $p = .038$; $d = -0.468$). (All other $ps > .10$).

Ratings of vignettes featuring the “no minority group” samples (men, White people, and right-handed people) in the car domain did not differ significantly from each other (all $ps = 1.00$). Figure 4 illustrates the results obtained in the “domain of car design”.

Research Domain of Education

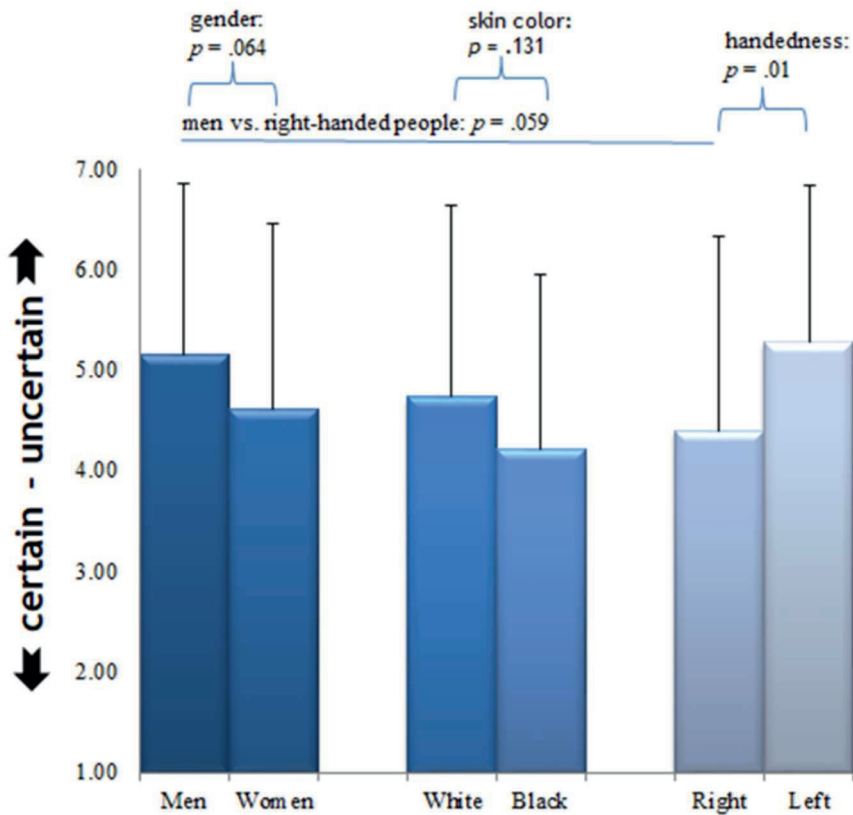


Figure 3. Results obtained in the “domain of education” were that participants expressed higher certainty (that a decision was a good one) when the introduction of a new method to measure intelligence for college admissions was based on tests with samples only women, compared to samples of only men. Further, certainty (that a decision was a good one) was higher with samples of only right-handed people, compared to samples of only left-handed people. Error bars show standard deviations.

Research Domain of Car Design

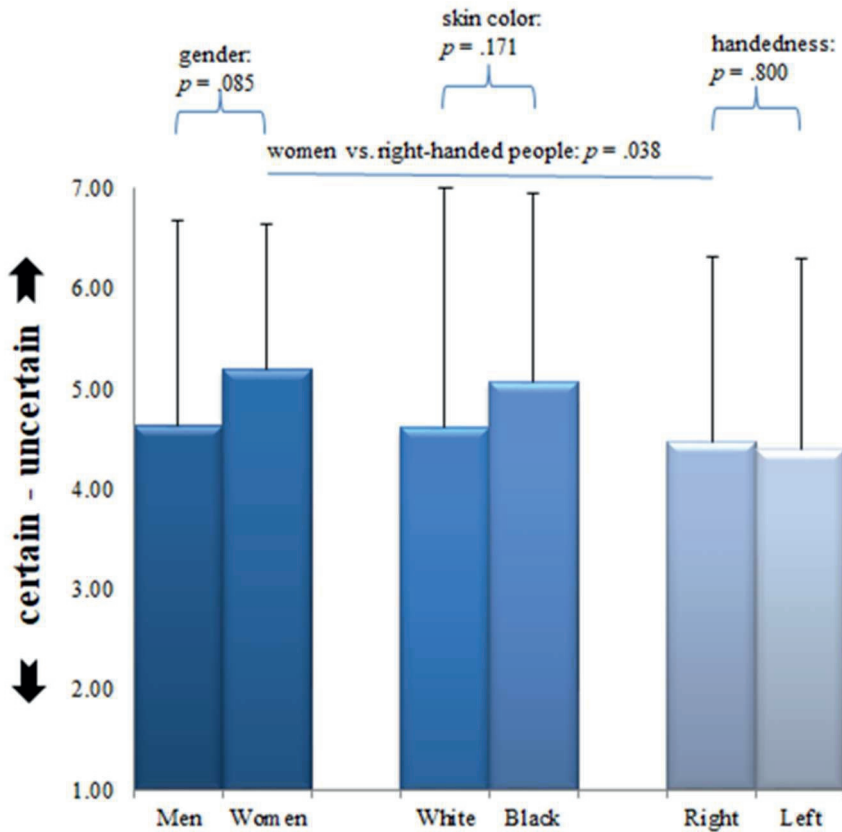


Figure 4. Results obtained in the “domain of car design” were that participants expressed higher certainty (that a decision was a good one) when the evaluation of a new car design was done by samples of men compared to samples of women. Further, certainty (that a decision was a good one) was higher with samples of men, compared to samples of left-handed people. Error bars show standard deviations.

Personal features and non-cognitive values in the assessment of inductive risk

In the next step, we evaluated whether and how our participants' personal features, in particular: sex, handedness, and race^{see footnote 66} influenced their assessment of inductive risk. Further, we evaluated whether and how a number of non-cognitive values such as "political interest" as well as "political", "race", and "gender attitudes" (in terms of "liberal", "moderate", and "conservative"), influenced the assessment of inductive risk in our experiment. To that end, we calculated analyses with the factors reported above, and including "sex", "race", "handedness", "political interest", "political attitude", "race attitude", and "gender attitude" (as reported by our participants) as covariates.

- Sex: The additional analysis, including the biological sex as a covariate revealed that the biological sex of our participants (male vs. female)⁷⁴ did not further modulate the effects obtained in the error type I-experiment (all $ps > .10$).

- Handedness: The additional analysis, including the handedness of our participants as a covariate, revealed no effects (all $ps > .05$).

- Race: The additional analysis, including race⁷⁵ as a covariate, revealed a main effect of "race", $F(4,84) = 7.372$; $p = .001$; $\eta^2_{\text{part}} = .028$, and an interaction of Minority Group $\times$ Race, $F(4,84) = 6.179$; $p = .001$; $\eta^2_{\text{part}} = .023$.

Post-hoc tests show that White participants expressed higher overall uncertainty (that a decision was a good one) in their ratings ($M = 4.87$; $SD = 1.80$), compared to Asian participants ($M = 4.30$; $SD = 1.91$; $t(967) = 4.524$; $p = .001$; $d = -0.31$).

Tests also show that Asian participants' ratings expressed significantly higher uncertainty (that a decision was a good one) with "no minority group" samples (men, White people, and right-handed people) ($M = 4.61$; $SD = 1.96$), as compared to "minority group" samples (women, Black people, and left-handed people) ($M = 4.02$; $SD = 1.83$; $F(1, 313) = 7.392$; $p = .007$; $\eta^2_{\text{part}} = .023$).

White and Black participants showed the opposite pattern of Asian participants: White participants expressed significantly higher uncertainty (that a decision was a good one) with "minority group" samples ($M = 5.11$; $SD = 1.71$) compared to "no minority group" samples ($M = 4.66$; $SD = 1.86$), $F(1, 652) = 10.614$; $p = .001$; $\eta^2_{\text{part}} = .016$. Black participants also expressed higher uncertainty (that a decision was a good one) with "minority group" samples ($M = 4.97$; $SD = 1.86$), compared to "no minority group" samples ($M = 4.13$; $SD = 1.91$), $F(1,61) = 3.098$; $p = .083$; $\eta^2_{\text{part}} = .048$. All other $ps > .15$.

⁷⁴ Ratings of subjects who reported to be gender "diverse" (here $n = 3$) were excluded from this analysis, due to a too small number or "diverse" subjects, to be taken into account as a sample "on their own".

⁷⁵ Races included were (Asian, Black/African Black, Hispanic/Latino, White, and "other").

The ratings of vignettes featuring the “no minority group” samples did not differ significantly between participants of different races (all p s > .20). Figure 5 depicts the results.

How “Race” modulates assessment of inductive risk

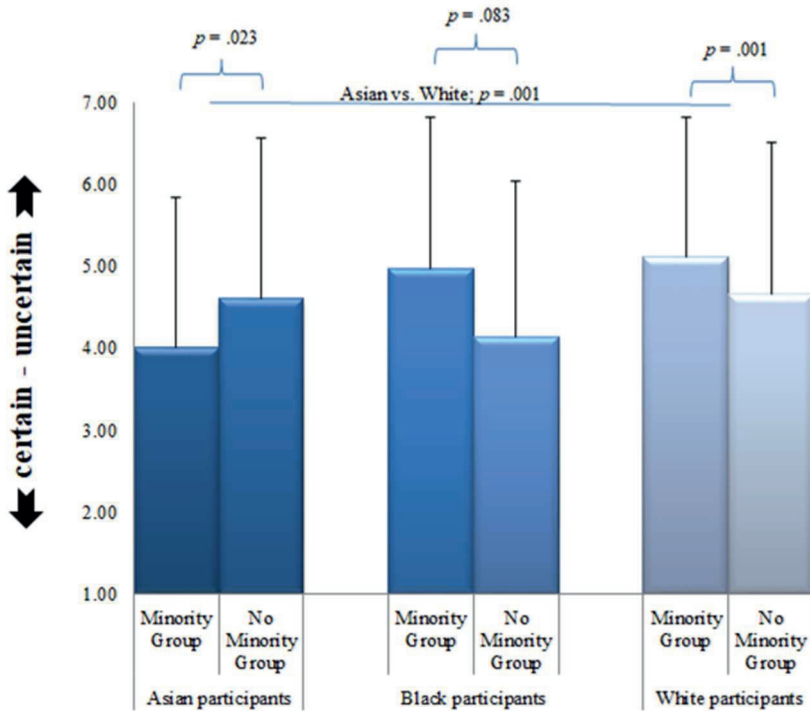


Figure 5. White participants expressed significantly higher overall uncertainty (that a decision was a good one) compared to Asian participants. Asian participants expressed less uncertainty (that a decision was a good one) with “minority group” samples (women, Black people, and left-handed people) compared to “no minority group” samples (men, White people, and right-handed people). Black and White participants showed the opposite “pattern”: Black and White participants expressed higher uncertainty (that a decision was a good one) with “minority group” samples compared to “no minority group” samples.

- Political interest: The additional analysis, including the “political interest” as a covariate revealed no significant effects (all $ps > .05$).

- Political Attitude: The additional analysis, including the “political attitude” of our participants as a covariate revealed a significant main effect of Political Attitude, $F(2,53) = 10.887$; $p < .001$; $\eta^2_{\text{part}} = .020$, a significant interaction of Minority Group $\times$ Political Attitude, $F(2,53) = 6.192$; $p = .002$; $\eta^2_{\text{part}} = .011$, a significant interaction of Sample $\times$ Political Attitude, $F(4,53) = 3.263$; $p = .011$; $\eta^2_{\text{part}} = .012$, and a significant interaction of Minority Group $\times$ Sample $\times$ Domain $\times$ Political Attitude, $F(8,53) = 2.004$; $p = .043$; $\eta^2_{\text{part}} = .015$. Remaining $p > .05$.

Post-hoc tests (Bonferroni-corrected) revealed that the political attitude of our participants influenced their assessment of inductive risk in the domains “healthcare” (see Figure 6) and “education” (see Figure 7), but not in the domain of “car design”.

In the *Domain of Health*, with samples of White people (i.e. “no minority group”, skin color sample) ($M = 5.48$; $SD = 1.692$), participants with “moderate” political attitude expressed significantly higher uncertainty (that the decision was a good one) compared to participants with “conservative” political attitude ($M = 3.50$; $SD = 2.15$), $t(31) = 2.955$; $p = .014$; $d = 1.06$. Comparisons between ratings of participants with “conservative” and “liberal” political attitude ($M = 4.82$; $SD = 1.95$) and comparisons between ratings of “moderate” vs. “liberal” participants were non-significant ($ps > .10$).

A similar pattern within the *Domain of Health* was found with samples of women (i.e. “minority group”, gender sample), ($M = 5.93$; $SD = 1.215$), where participants with “liberal” political attitude expressed higher uncertainty (that the decision was a good one) compared to participants with “conservative” political attitude ($M = 3.25$; $SD = 1.603$) ($t(27) = 4.962$; $p < .001$; $d = 2.002$), and compared to participants with “moderate” political attitude ($M = 4.00$; $SD = 2.00$) ($t(33) = 4.0125$; $p = .001$; $d = 1.242$). Participants with “conservative” and “moderate” political attitude did not differ significantly ($p = .627$, in the Domain of Health, regarding samples of women).

A further result within the *Domain of Health* was with samples of Black people, (i.e. “minority group, skin color sample), where the ratings of participants with “conservative” attitude expressed higher uncertainty (that the decision was a good one) ($M = 6.22$; $SD = 0.83$) compared to participants with “moderate” ($M = 4.06$; $SD = 1.78$) ($t(24) = 2.903$; $p = .015$; $d = -1.408$), and “liberal” political attitude ($M = 4.40$; $SD = 1.94$) ($t(50) = 2.7685$; $p = .022$; $d = -1.001$). Figure 6 depicts the effects of the political attitude of our participants in the *Domain of Health*.

How “Political Attitude” modulates assessment of inductive risk in the domain of healthcare

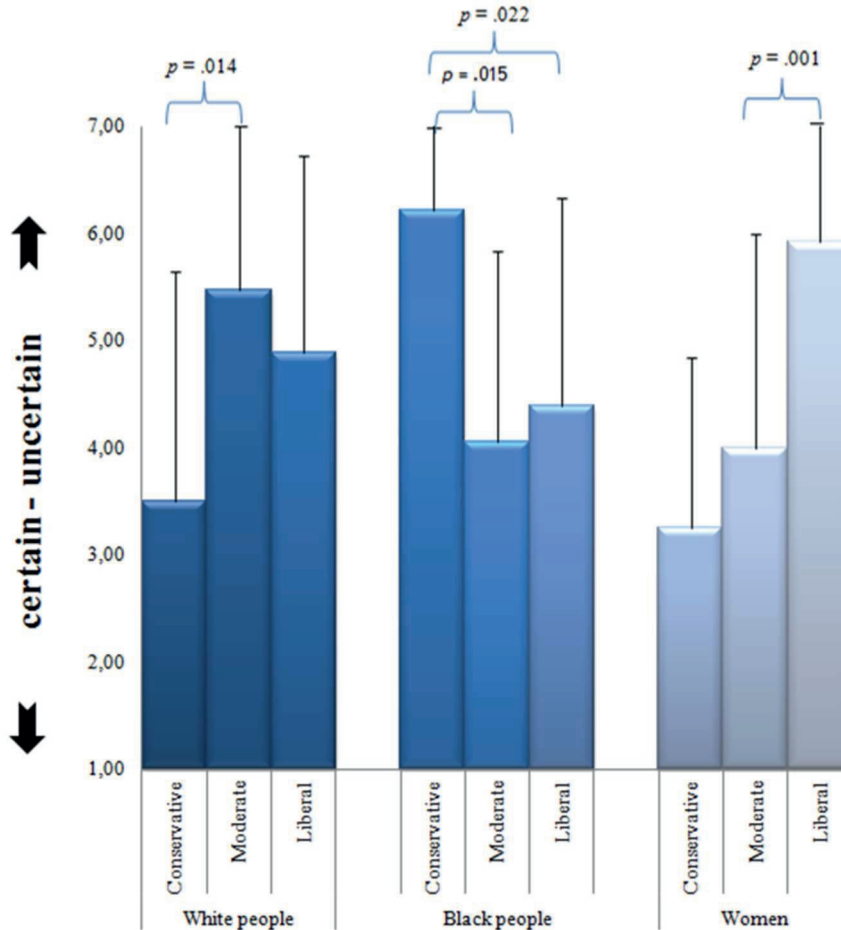


Figure 6. The figure depicts the results of the influence of political attitude on the assessment of inductive risk within the “domain of health”. With samples of White people, “conservative” participants, compared to “moderate” participants, expressed significantly higher certainty (that the decision was a good one). With samples of Black people, “conservative” participants also expressed significantly higher uncertainty compared to “moderate”, and compared to “liberal” participants. With samples of women, “moderate” participants expressed higher uncertainty compared to “liberal” participants. Error bars indicate standard deviations.

In the *Domain of Education*, post-hoc tests show that with samples of left-handed people (i.e. “minority group”, handedness sample) participants with “liberal” attitude expressed higher uncertainty (that the decision was a good one) ($M = 5.61$; $SD = 1.22$), compared to participants with “moderate” political attitude ($M = 4.36$; $SD = 2.06$) ($t(45) = 2.610$; $p = .035$; $d = 0.826$).

Comparisons of ratings between “conservative” participants ($M = 5.44$; $SD = 1.424$) and “liberal” as well as “conservative” and “moderate” participants were non-significant ($ps < .25$). Figure 7 depicts the result of political attitude in the *Domain of Education*.

How “Political Attitude” modulates assessment of inductive risk in the domain of education

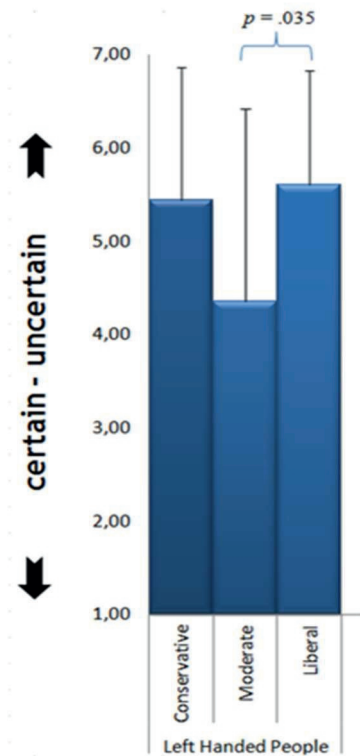


Figure 7. The figure depicts the results of the influence of political attitude on the assessment of inductive risk within the “domain of education”. With samples of left-handed people, “moderate” participants, compared to “liberal” participants, expressed significantly higher uncertainty (that the decision was a good one). Error bars indicate standard deviations.

- Race Attitude: The additional analysis, including the “attitude towards race related social issues” (in the following referred to as “Race Attitude”) ratings⁷⁶ of our participants as a covariate revealed a significant main effect of Race Attitude, $F(2,53) = 7.791$; $p = .001$; $\eta^2_{\text{part}} = .014$, and a significant interaction of Minority Group $\times$ Race Attitude, $F(2,53) = 4.211$; $p = .015$; $\eta^2_{\text{part}} = .008$.

Post-hoc tests (Bonferroni-corrected) showed that participants who referred to themselves as “conservative” expressed – in general – lower uncertainty (that the decision was a good one) in their assessments of inductive risk of error type I ($M = 4.35$; $SD = 1.97$) compared to participants who referred to themselves as having a “liberal” race attitude ($M = 4.87$; $SD = 1.78$; $t(859) = 3.881$; $p < .001$; $d = 0.283$). Comparisons between ratings of “conservative” and “moderate” ($M = 4.56$; $SD = 1.851$) and between “liberal” and “moderate” participants were non-significant (all $ps > .050$).⁷⁷

Participant with “moderate” race attitude expressed significantly higher uncertainty with *no* minority groups (men, White people, right-handed people) ($M = 4.81$; $SD = 1.757$) compared to minority groups (women, Black people, left-handed people) ($M = 4.29$; $SD = 1.914$), $F(1,259) = 5.154$; $p = .024$; $\eta^2_{\text{part}} = .020$. Participants with “liberal” race attitude expressed the opposite, their uncertainty was significantly higher with minority groups ($M = 5.02$; $SD = 1.726$), compared to *no* minority groups ($M = 4.73$; $SD = 1.814$), $F(1,604) = 3.974$; $p = .047$; $\eta^2_{\text{part}} = .007$. Participants with “conservative” race attitude ratings of minority ($M = 4.35$; $SD = 1.972$) vs. *no* minority groups ($M = 4.24$; $SD = 2.102$) did not differ significantly ($p = .357$). Figure 8 depicts the effects of the race attitude of our participants on inductive risk of error type I assessment.

- Gender Attitude: The additional analysis, including the “attitude towards gender related social issues” (in the following referred to as “Gender Attitude”) ratings of our participants as a covariate revealed a main effect of Gender Attitude, $F(2,53) = 8.208$; $p < .001$; $\eta^2_{\text{part}} = .015$.

Post-hoc tests (Bonferroni-corrected) show that participants with “liberal” gender attitude expressed significantly higher uncertainty (that the decision was a good one) by their ratings ($M = 4.88$; $SD = 1.780$), compared to participants with “moderate” gender attitude ($M = 4.49$, $SD = 1.902$) ($t(907) = 2.955$; $p = .010$, $d = -0.215$) and compared to participants with “conservative” gender attitude ($M = 4.36$, $SD = 1.924$) ($t(838) = 3.64286$; $p = .001$, $d = -0.286$). The comparison between the ratings of “conservative” and “moderate” participants was non-significant ($p = 1.00$). Figure 9 shows the effects of the gender attitude of our participants on the assessment of error type I inductive risk.

⁷⁶ Towards “race related social issues, of the $N = 374$ participants, $n = 85$ referred to themselves as “conservative”, $n = 87$ as “moderate”, and $n = 202$ as “liberal”.

⁷⁷ Although the ratings of “moderate” compared to “liberal” participants differed marginally significantly ($p = .054$).

How “Race Attitude” modulates assessment of inductive risk

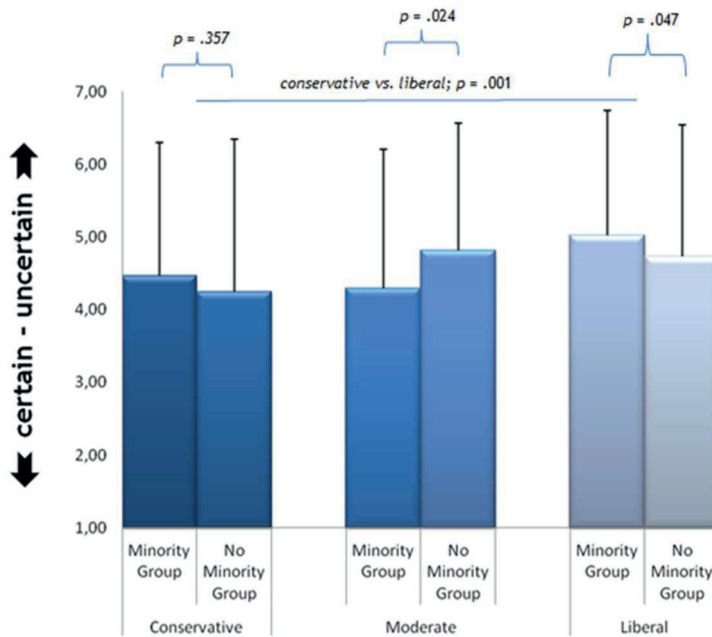


Figure 8. The figure depicts the effects of the race attitude of our participants on inductive risk of error type I assessment. The participants with “conservative” race attitude expressed significantly lower uncertainty (that the decision was a good one), compared to the participants with “liberal” race attitude. The participants with “moderate” race attitude were more uncertain (that the decision was a good one) with “no minority group” samples (men, White people, and right-handed people), compared to “minority group” samples (women, Black people, left-handed people), while participants with “liberal” race attitude were more uncertain (that a decision was a good one) with “minority group” samples, compared to “no minority group” samples.

How “Gender Attitude” modulates assessment of inductive risk

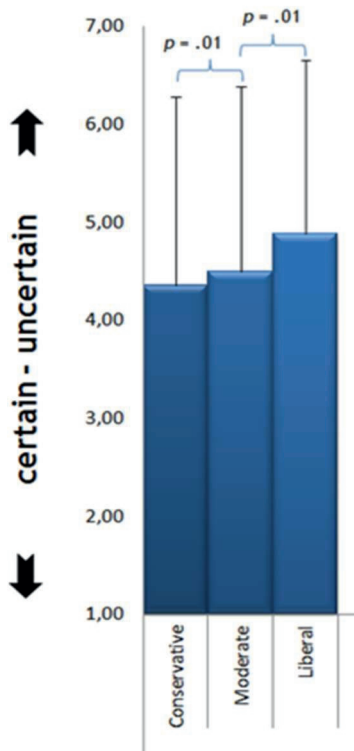


Figure 9. The figure depicts the results of the influence of gender attitude on the assessment of error type I inductive risk. The participants with “liberal” gender attitude expressed significantly higher uncertainty (that the decision was a good one) with compared to the participants with “moderate” and “conservative” gender attitude.

3.2 Experiment: Error Type II

The design, material, and procedure of the experiment on the assessment of the inductive risk of committing an error type II resembled the error type I experiment, with the exception that the third sentences of the vignettes of the error type II experiment reported the decision to *maintain* the situation as it is, namely, the decision of not introducing a new medicine, intelligence measure or car design into the market or college admission tests. The structure of the vignettes was the same as in error type I experiment with the third sentence as follows:

[Third sentences of error type II-experiment's vignettes]:

"The decision was made to withhold [the new medicine from the market⁷⁸ / the new method to measure intelligence from examinations for college admissions⁷⁹ / the new car from the market⁸⁰].

This phrasing was chosen to assess inductive risk in (error type II) cases where the risk is that the null hypothesis is maintained although it is not true, and should have been rejected.

Participants

A total N = 354 participants took part in the Error Type II experiment, mean age 33.43 years, *SD* = 9.14 (age unknown for n = 6 participants), n= 122 female, n = 231 male, n = 1 diverse⁸¹. The participants received a small monetary payment. Each participant rated three vignettes, one for each of the three domains of research we used (i.e., healthcare, education, and car design). Approximately half of the participants (n = 174) were presented with and judged studies including samples composed of only "minorities" (i.e., women, Black people, and left-handed people). The other half of the participants (n = 180) was presented with and judged studies including samples composed of only "no minorities" (i.e. men, White people, and right-handed people). The combinations of research domain and sample were counterbalanced across participants to meet the criteria of a complete design, analogous to the design of Error Type I Experiment (see figure 1 above).

Results

⁷⁸ Domain of healthcare.

⁷⁹ Domain of education.

⁸⁰ Domain of car design.

⁸¹ Of the participants, n = 166 reported to be "White", n = 21 "Black", n = 131 of Asian, n = 18 of Hispanic (or Latino), and n = 18 of other ethnical origin. Three Hundred of the participants reported to be right-handed, n = 42 left-handed, and n = 12 ambidextrous.

The analysis, including the same factors as reported for experiment 1 (and in line with the factorial design of the experiment) revealed that none of the factors influenced the ratings (all $ps > .15$).

However, the assessment of risks in the Error Type II scenarios was influenced by some of the non-cognitive values of our participants (see results below). It was not influenced by personal features such as race and handedness ($p > .10$). The “sex” of our participants modulated the ratings slightly, in that way that in the Error Type II scenarios, female participants expressed - in general - marginally significant higher uncertainty ($M = 3.91$; $SD = 1.84$), compared to male participants ($M = 3.69$; $SD = 1.80$; $F(1, 1057) = 3.573$; $p = .059$; (see figure 10)⁸².

The additional analysis, including Political Interest, Political Attitude, and Gender Attitude as a covariate revealed no significant effects (all $ps > .10$). However, Race Attitude modulated the ratings, as follows.

The additional analysis, including the “attitude towards race related social issues” (in the following referred to as “Race Attitude”) ratings⁸³ of our participants as a covariate revealed a significant interaction of Minority Group $\times$ Race Attitude, $F(2,53) = 5.187$; $p = .006$; $\eta^2_{\text{part}} = .010$. All other $ps > .10$.

Post-hoc tests (Bonferroni-corrected) showed that participants with “conservative” race attitude expressed significantly higher uncertainty with *no* minority groups (men, White people, right-handed people) ($M = 4.05$; $SD = 1.900$) compared to minority groups (women, Black people, left-handed people) ($M = 3.58$; $SD = 1.761$), $F(1,265) = 4.285$; $p = .039$; $\eta^2_{\text{part}} = .016$. Participants with “liberal” race attitude expressed the opposite (similar result as with Error Type I scenarios), their uncertainty was significantly higher with minority groups ($M = 3.91$; $SD = 1.714$), compared to *no* minority groups ($M = 3.57$; $SD = 1.876$), $F(1,541) = 4.739$; $p = .030$; $\eta^2_{\text{part}} = .009$. Participants with “moderate” race attitude’s ratings of minority ($M = 3.96$; $SD = 1.759$) vs. *no* minority groups ($M = 3.73$; $SD = 1.874$) did not differ significantly ($p = .303$). Figure 11 depicts the effects of the race attitude of our participants on inductive risk of error type II assessment.

⁸² The ratings of subjects who reported to be gender “diverse” (here only $n = 1$) were excluded from this analysis, due to a too small number of “diverse” subjects to be taken into account as a sample “on their own” of diverse participants.

⁸³ Regarding “race related social issues”, of the $N = 354$ participants, $n = 89$ referred to themselves as “conservative”, $n = 84$ as “moderate”, and $n = 181$ as “liberal”.

The personal feature of gender in the assessment of inductive risk of error type II

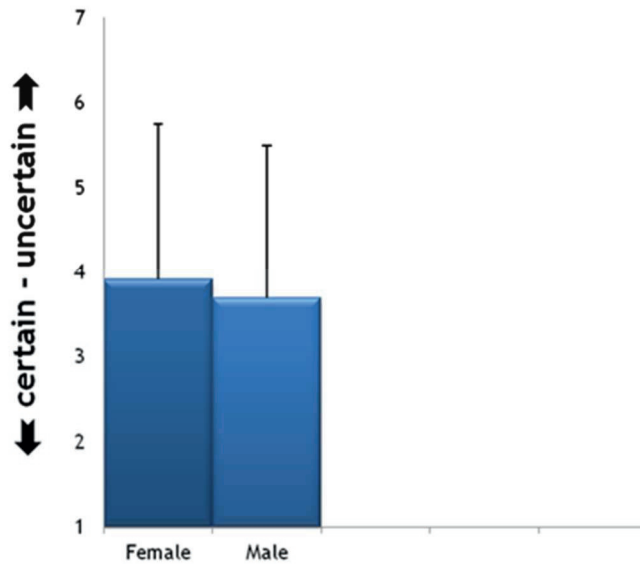


Figure 10. Female participants compared to male participants expressed higher uncertainty (that the decision was a good one) in the assessment of error type II scenarios.

How race attitude modulates the assessment of error type II inductive risk

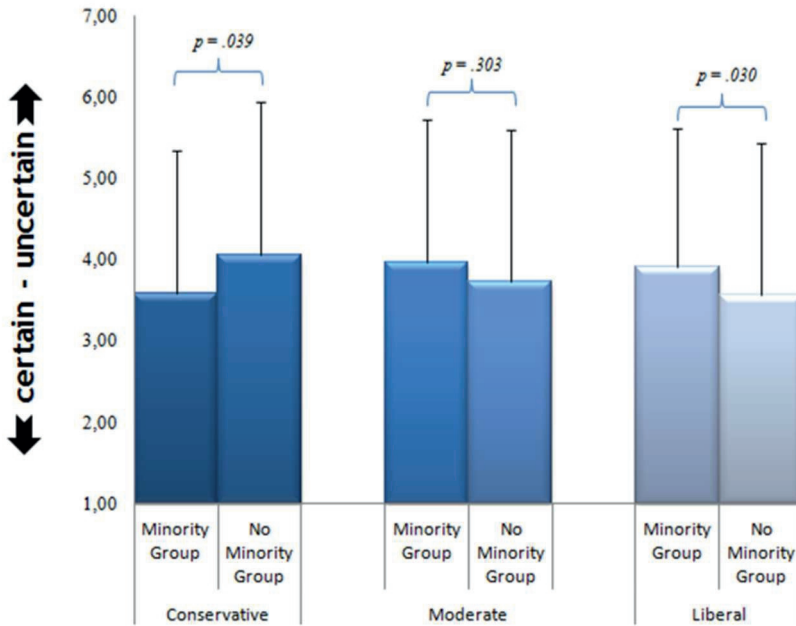


Figure 11. The figure depicts the results of the influence of race attitude of our participants on the inductive risk of error type II scenarios. The participants with “conservative” race attitude expressed higher uncertainty (that the decision was a good one) with “no minority” group samples compared to “minority” group samples, while participants with “liberal” race attitude expressed higher certainty with the “no minority” group samples compared to the “minority” group samples.

The comparison between assessment of inductive risk of error type I and error type II

Finally, the comparison between Error Type I and Error Type II ratings revealed an asymmetry of certainty. Assessment of inductive risk in error type I-scenarios was associated with higher uncertainty (that the decision was a good one) ($M = 4.68$; $SD = 1.851$) compared to inductive risk in error type II-scenarios ($M = 3.76$; $SD = 1.818$), $F(1, 2182) = 138.261$; $p = .001$; (see figure 12).

The result suggests higher uncertainty (that the decision was a good one) associated with changes (by introducing a certain novelty, as it was the case in our error type I-scenarios) compared to higher certainty (that the decision was a good one) associated with the maintenance of the situation as it is (by *not* introducing a certain novelty, as it was the case in our error type II-scenarios).

Comparison of the assessment of inductive risk in error type I and error type II scenarios

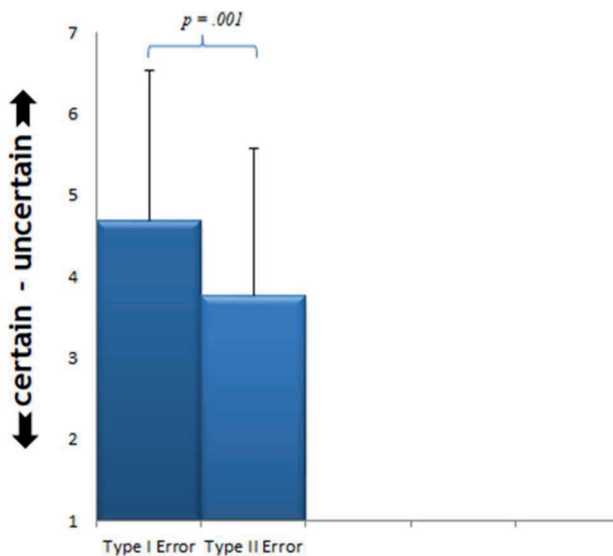


Figure 12. Participants who assessed error type II scenarios expressed higher certainty (that the decision was a good one) compared to participants who assessed error type I scenarios.

4 Discussion. The psychology of inductive risk

In this study, we explored how lay people judge possible cases of inductive risk. In particular, we were interested in understanding whether people's confidence in their judgments varies depending on the domain in which an error may be committed (Q1), people's personal features (Q2), and the possible types of error (Q3).

Our results show that various factors determine how people react to inductive risks. People's level of confidence was not determined by the domain in which a possible mistake was made, but by the combination of domain and group affected by the consequences of a possible error (Q1). Participants expressed similar levels of uncertainty across domains (neither certain - nor uncertain / slightly uncertain). However, which group triggered their uncertainty varies across areas (e.g. men in the healthcare domain, women in the product design domain). Our results show that being potentially harmed by the consequences of a mistaken decision did not predict how people react to errors (Q2). However, some personal features had a role in determining people's reactions. The results of the first experiment revealed a main effect of "race", e.g. White participants expressed overall more uncertainty than Asian participants. Moreover, our results show that participants endorsing different political values expressed different levels of uncertainty. Finally, an asymmetry in the reactions to possible false positives and possible false negatives was found (Q3). While participants expressed uncertainty when evaluating decisions taken on the basis of possible false positives, they were generally confident that the decisions taken in the possible false negative scenarios were the right decisions to take.

This asymmetry in reactions can be explained as the result of the *status quo bias*. Several studies in psychology show that when people face decision problems, such as the selection of health plans and retirement investments, and both the options of retaining the status quo (maintaining the situation as it is by selecting last year's health plan) and the option of changing it (e.g. changing the situation by selecting a new health plan) are available, they tend to make decisions that maintain the current situation (Samuelson & Zeckhauser 1988).

The status quo bias is an implication of what is called *loss aversion* in the psychological literature (Kahneman & Tversky 1979; Kahneman et al. 1991). Psychological studies show that the "aggravation that [people] experiences in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount" (Kahneman & Tversky 1979, 279). People would thus generally find avoiding losses more pleasing and desirable than acquiring gains.

In our study, a decision based on a false positive meant (mistakenly) changing the status quo (e.g. changing the existing situation by introducing a novelty into the market), while a decision based on a false negative meant (mistakenly) maintaining the status quo. Participants' lower levels of uncertainty in the assessment of possible mistaken decisions taken on the basis of false negatives may be seen as an effect of these mechanisms:

maintaining the status quo in cases of inductive risk would be a way to avoid the possible losses and aggravation that one may experience when (mistakenly) changing a situation by introducing a new medicine, intelligence measure or car design.

Douglas has argued that in cases of inductive risk which errors are seen as more tolerable depends on the non-epistemic consequences they involve (Douglas 2000; see section 2 of this chapter). The fact that our participants expressed lower levels of uncertainty in the possible false negatives scenarios might mean that overall they perceived the possible consequences of false negatives as more (morally, socially or economically) tolerable than the possible consequences of false positives.

Martie Haselton and David Buss have provided an adaptationist explanation for human beings' "preference" for false negatives (Haselton & Buss 2000; Haselton & Buss 2003). It is likely, Haselton and Buss argue, that human beings have evolved mechanisms that make them prone to incur false negatives and to avoid false positives. Their idea is then that the preference for false negatives may have been evolved to solve recurrent adaptive problems in EEA, such as estimating the risk of aggression or food poisoning (Haselton & Buss 2003, 34-38). Our study does not provide evidence relevant to understanding whether the asymmetry of reactions to possible errors has an evolutionary origin (see section 3.2 of Chapter 2). However, Haselton and Buss' theory provides a framework to explain the asymmetry of reactions to errors that future research might investigate further.

The present study focuses on the descriptive side of inductive risk. However, besides extending our understanding of how people react to uncertainty in science, these results are relevant to the discussion of normative questions in philosophy of science, too.

First of all, our results are relevant to discuss the value-free ideal of science (see section 1 of Chapter 1). Our study shows that some personal features (e.g. race) and political attitudes influence the way people judge possible cases of inductive risks. For instance, when assessing possible type I errors in the healthcare domain, conservatives expressed higher uncertainty compared to moderates and liberals when the sample only included Black people, whereas liberals expressed more uncertainty than the other participants when the sample only included women (see table 6). The image of science provided by the value-free ideal, in which preserving the epistemic authority of science means excluding non-cognitive values from the decisions concerning evidence and scientific hypotheses, might then be *unattainable*. If non-cognitive values influence how people judge possible cases of inductive risks, and people are systematically more sensitive to the consequences of false positives than to the consequences of false negatives, the very possibility of having value-free assessments of scientific hypotheses might be questioned.

A limitation of our study is that it focuses on lay people's judgments and not on scientists' judgments. Thus, our results may be seen as unrepresentative of how scientists judge cases of inductive risks and the claim that these results challenge the value-free ideal may be questioned. At the end of the day, scientists (and political institutions), and not lay

people, make decisions in cases of inductive risks, and they are supposed to have a better understanding of the reliability of scientific methodologies and chances of committing errors than non-experts do. However, research in psychology has shown that values and prior beliefs influence how both lay people and scientists assess the quality of evidence and the reliability of methodologies, and the way they estimate the risks and benefits of scientific practices (e.g. Koehler 1993; Kahan et al. 2010; see section 2 of this chapter).

More research is needed to clarify to what extent and in what respects lay people's and scientists' reasoning in cases of inductive risk differ (see for instance Slimak & Dietz 2006, 1698-1699). However, clarifying how non-cognitive values influence lay people's judgments of possible cases of inductive risk is relevant to discuss some questions in the Science & Values debate.

Governmental institutions have recently developed international projects aiming at the integration of citizens' values into scientific research, such as the "Science with and for Society" project in the EU's Horizon 2020 research and innovation programme. This project adopts the "Responsible Research and Innovation" approach, which requires that all societal actors (e.g. scientists, citizens, and policy makers) collaborate in scientific research and innovation agendas in order to align them "with the values, needs and expectations of European society" (Figueiredo Nascimento et al. 2016, 12). The idea underlying projects like this one is that scientific research should focus on "topics that really matter for society and are closer to real needs" (Figueiredo Nascimento et al. 2016, 28), and that it should be conducted in a way that exemplifies and promotes the needs and values of the society.

Studies like the one presented in this chapter can provide the basis to understand how citizens judge possible cases of inductive risks and to consistently consider their values in scientific decisions. However, a question that may be asked is to what extent it is *desirable* to pursue a citizens'-values-laden science of this kind.

Some philosophers of science have defended the inclusion of citizens' values and needs in science. They have argued that "scientists should privilege the [non-cognitive] values endorsed by the public or its representatives" in cases of inductive risk (Schroeder 2017, 1045), because the decisions taken in these phases can have various social, economic, and ethical consequences that may affect citizens (e.g. introducing a new medicine into the market that may harm some groups of the population). It is not acceptable, these philosophers argue, that these decisions are taken by "a minority elite" (i.e. scientists, Douglas 2005, 156). In these decisions, the voices of citizens should be heard. Moreover, some philosophers have argued that since taxpayers provide the funding for research, they should have a say in how money is spent and scientists have no right to refuse public discussion of scientific research (Feyerabend 1978; Nature 2004; Douglas 2005; Reddy 2009).

Some examples from scientific practice have shown that including citizens' values and concerns can have a positive influence on science, in particular in estimating the possible risks that the use of certain technologies may bring about. For instance, Douglas argues that

the joint work of government agencies, scientists, citizens, and members of oil companies led to a better understanding of the problems involved in marine oil trade in Valdez, Alaska (Douglas 2005, 159). Specifically, the joint work of these groups made it possible to understand the environmental risks of marine oil trade better and to develop techniques able to avoid oil spill that marine oil trade could cause in that area (ibid.).

However, while there are ethical and social reasons for including the values, needs, and expectations of citizens in scientific research, this inclusion may be, in certain cases, undesirable.

To illustrate, take the medical vignette as an example. If protecting the health of some groups (e.g. White people) is considered by some citizens as more valuable than protecting the health of other groups (e.g. Black people; see figure 6 on political attitude and healthcare domain in the previous section), a possible consequence of privileging citizens' values is that clinical research might overlook differences in frequency of diseases and reactions to drugs among people with different ethnic backgrounds, and that potentially dangerous medicines are introduced into the market. Consider also our results on people's reactions to different types of error. The present study has shown that our participants prefer maintaining the status quo and seem more willing to tolerate the consequences of false negatives than the consequences of false positives; should scientists do the same in cases of uncertainty? False negatives may have undesirable consequences, just like false positives. For instance, deciding that a medicine will not be introduced into the market when it is actually a safe medicine may have the consequence of limiting people's access to cures that may be critical to their survival. Favouring false negatives over false positives may then not be the right choice in some scenarios and following people's "preference" for false negatives may lead to undesirable consequences.

These examples show that aligning the decisions taken in cases of inductive risk with citizens' judgments may, in some cases, lead to epistemic failures, i.e., to a deficient understanding of phenomena and a mistaken assessment of hypotheses (e.g. accepting the hypothesis that a medicine is safe when it is actually unsafe), and to serious non-epistemic consequences for the society (e.g. harming people's health). The citizens'-values-laden image of science defended by institutions and some philosophers of science may then be often undesirable.

It is an open question how citizens' values can be integrated into scientific research so as to meet the needs of citizens and avoid the problems presented in this section. Schroeder has argued that citizens' values should be seriously considered by scientists only when "informed and substantially reasonable" (Schroeder 2017, 1045). However, in many cases it may be unclear what should count as informed and reasonable values or which values should prevail among different kinds of informed and reasonable values when a decision has to be made. One way to deal with this issue may be considering the aforementioned epistemic and social consequences of making a decision on the basis of the values endorsed by citizens.

When these values lead to epistemic failures and/or to make a moral wrong (e.g. harming people's health), scientists may have good reasons for refusing the integration of these values into scientific research.

5 Conclusion

Inductive risk challenges the value-free ideal of science by revealing the beneficial and fundamental role of non-cognitive values in science. In cases of uncertainty, decisions over scientific hypotheses and methodologies legitimately involve considering ethical, social, or economic factors. The study presented in this chapter contributes to this discussion in two fashions. First, it provides empirical evidence relevant to understand the psychology of inductive risk, namely information on how people judge possible cases of inductive risks and what roles the domains in which errors may be committed, personal features, political values, and types of error play in these judgments. Second, it provides normatively relevant information to discuss various philosophical questions over the role of values in science, such as questions over the value-free ideal and the importance of citizens' values in science.

Before going further, a final remark. The psychological literature reports mixed results on the effects of communicating scientific uncertainty to citizens. Some show that it leads to fatalism among citizens, that it tends to make lay people scared and outraged (e.g. Johnson & Slovic 1998) and that people see scientific uncertainty as a signal of the incompetence of scientists (Johnson & Slovic 1995). In contrast, other studies report that the communication of scientific uncertainty increases people's interest in science and technologies and does not reduce their trust in scientists (Retzbach & Maier 2015, 446). To date, research has not provided a clear, coherent picture of what determines people's reactions to scientific uncertainty: "there is no clear pattern to report at the moment, just a series of studies that seem to raise more questions than they answer" (Jensen et al. 2017, 47). By providing information on what triggers uncertainty in people (e.g. type of error) and how personal features (e.g. race, political values) predict how people judge possible cases of inductive risk, the results of the present study may also be relevant to find strategies to communicate scientific uncertainty to the public able to avoid mistrust, panic, and fatalism.

Chapter 6

Conclusion

A large body of literature has investigated what influence values have and should have on science, but several questions remain to be analysed. The central chapters of this dissertation address some of these questions. In this chapter, I conclude by summarizing the claims I have defended in this dissertation and by introducing some future lines of research.

We need clear explications and strategies to use cognitive values in theory appraisal

The philosophical literature provides vague explications for some of the cognitive values included in the traditional lists. The importance of these values for theory appraisal is often taken for granted and it is not clear how they can be used in theory appraisal. In other words, it is not clear which aspects of a theory or program should be considered to assess whether that certain theory or program satisfies a certain value.

Vague explications and unrefined strategies are of little use if we aim at clarifying the roles of values in science, and to make sure that employing cognitive values in theory appraisal makes us able to appreciate the virtues of scientific theories or research programs. In the second chapter, I have started filling this gap by analysing one of the values that have attracted little attention in the philosophical literature, namely fruitfulness. I have explicated fruitfulness as the ability of programs to extend their content and suggested considering research questions and discovery heuristics to assess this ability. Moreover, I have used my account to assess the fruitfulness of Evolutionary Psychology.

Focusing on research questions and discovery heuristics revealed to be a successful strategy to highlight and discuss some problems in the methodological assumptions of Evolutionary Psychology, in particular adaptationism. However, the assessment of the fruitfulness of theories and programs may involve also other tools, and future research may develop the analysis of fruitfulness further by considering these other tools, such as the methodologies used to test hypotheses and predictions. In addition, case studies from different fields may be developed, to understand whether the same tools are relevant to assess the fruitfulness of programs in different domains.

One of the contributions of the second chapter to the Science & Values debate is highlighting the importance of having clear explications and guidelines to use cognitive values in theory appraisal. Fruitfulness is just one of the values that have received little attention in the philosophical literature. Future research may develop similar analyses for other cognitive values, such as scope and external consistency.

Contextual factors have a role in determining whether and how scientific theories satisfy cognitive values and the desirability of cognitive values

The philosophical literature explains the importance of values like fruitfulness and explanatory power in terms of epistemic or pragmatic desirability: values are desirable because they can provide scientists with epistemic and/or pragmatic reasons to prefer a theory over its alternatives. However, there is disagreement on which values are epistemically and/or pragmatically desirable. In Chapter 3, I have used the Evolutionary Theory of stalking (Duntley & Buss 2012) as a case study to analyse the desirability of explanatory power, external consistency, and fruitfulness. On the basis of this case study, I have defended three main theses. First, I have argued for the situatedness of theory appraisal: the assessment of scientific theories is always done in a context, which is historically and culturally situated. This means that some local contingencies, such as the methodological assumptions endorsed by the scientific community and the availability of relevant evidence, may facilitate (or obstruct) the chances to appreciate whether and to which degree a theory satisfies certain cognitive values.

Second, I have argued for the contextual desirability of values: whether values are desirable for epistemic or pragmatic reasons may depend on some contextual factors. For instance, the case study shows that the availability of relevant evidence and the reliability of the background theories determine whether explanatory power and external consistency are desirable for epistemic reasons. Future research may expand this analysis by specifying other contextual factors, besides evidence's availability and background theories' reliability, that may influence the desirability of values, such as the influence of social, ethical, and political assumptions on the scientific community.

Third, I have argued for the thesis of the interrelationship between values: in order to understand whether a theory satisfies a certain value, it is important to consider how and whether the theory satisfies other values. This is because how and whether a theory satisfies certain values may facilitate (or not) the fulfilment of other values. To illustrate this point, I have clarified some of the relations between explanatory power, external consistency, and fruitfulness. However, values may be interrelated in manifold ways, which may be explored in future research. For instance, studies could investigate the relation between explanatory power and scope.

All in all, the study I have developed in the third chapter shows that when it comes to use cognitive values to assess scientific theories, we have to go local. That is, in order to achieve an accurate appraisal of theories, we have to consider specific factors of the context in which theories are assessed. My analysis has focused on Evolutionary Psychology. An open question is whether and how the field (e.g. psychology, biology, economics) makes any difference for the desirability of values, i.e., whether the desirability of values varies across fields. Although some philosophers have claimed that, depending on the field, different values would be relevant to theory appraisal, they have provided little evidence to justify this claim. Case studies from different fields should be developed to clarify this point.

Non-cognitive values can play a cognitive role in theory appraisal

Advocates of the value-free ideal argue that the influence of non-cognitive values on scientific reasoning threatens the epistemic authority of science. In the fourth chapter, I have challenged this view and argued that some non-cognitive values can play a cognitive role in science, i.e., they can be epistemically beneficial to the assessment of scientific theories. On the basis of a case study (the account of human mating in Evolutionary Psychology), I have argued that feminist values have positively contributed to theory appraisal in various ways, such as by raising sensitivity to evidence that was neglected because of gender bias.

This chapter also raises some questions over the traditional distinction between cognitive and non-cognitive values. If values are labelled as cognitive/non-cognitive depending on whether they can perform a certain function (i.e. fostering the attainment of knowledge), it seems that some of the values that are commonly included in the group of non-cognitive values, e.g. feminist values, should be listed among the cognitive ones.

I have also argued that not any non-cognitive values (e.g. racist values) can play the cognitive role feminist values play in my case study and that whether values are able to have this cognitive function depends on the different epistemic commitments they involve. Future research should explore this point further, for instance by investigating the epistemic benefits provided by other kinds of non-cognitive values in other fields (e.g. the social value of sustainability in economics and finance).

One way to explain the cognitive role played by feminist values in my case study is in terms of cognitive values: feminist values were epistemically beneficial because instrumental in improving the external consistency and predictive accuracy of theories. Thus, one possible line of future research is investigating how non-cognitive and cognitive values can be interrelated to each other, by focusing on which cognitive values social, political, economic, and ethical values are able to promote.

Non-cognitive values affect the way people judge possible cases of inductive risks

Critics of the value-free ideal appeal to inductive risk to demonstrate the legitimate roles non-cognitive values play in science. The fifth chapter of this dissertation takes an original approach to the argument from inductive risk by investigating the psychology of inductive risk with an experimental study. The results of this study show that various factors (e.g. the combination of the area in which an error may be committed and the group that may be harmed, the type of error, and the political values endorsed by participants) have a role in determining how people judge possible cases of inductive risk.

I have used the results of this study to discuss the attainability of the value-free ideal of science and whether citizens' values should be integrated into scientific research. However, our study raises several questions that need to be explored further: do experts' and non-experts' judgments about cases of inductive risks differ? What does it mean to *align* scientific research with citizens' values, needs, and expectations as requested by governmental institutions? Can integrating citizens' values into scientific research threaten the epistemic authority of science? Answering these questions is particularly important. It would make a significant contribution to psychology by clarifying whether and to what extent *expertise* makes a difference in how people judge and react to uncertainty in science, and in particular to cases of inductive risks. Moreover, answering these questions would also have clear practical implications, by providing information relevant to the development of policy measures aimed at engaging citizens into scientific research.

The fifth chapter of my dissertation highlights the importance of the experimental methods for the Science & Values debate. The survey we conducted made it possible to gather information that conceptual analysis could not offer (e.g. information over how people react to different types of error). To date, these methods have been rarely used to investigate questions over the roles of values in science, but future studies should fill this gap. For instance, future research may develop surveys aimed at investigating the influence of cognitive values on scientists' assessment of scientific theories. These surveys may help us to understand how scientists construe specific cognitive values like explanatory power and predictive accuracy, and to clarify whether different cognitive values are relevant to theory appraisal in different fields. This does not mean that the Science & Values literature should

abandon the case study method. Rather, what this dissertation shows is that combining case studies and experimental methods makes it possible to ask new questions and provide the appropriate evidence base to investigate these questions.

The arguments I have developed in my dissertation challenge the value-free ideal as a descriptively inaccurate and normatively undesirable ideal of science. They emphasize the importance of cognitive values for the assessment of scientific theories, they show that non-cognitive values play a role in scientific reasoning, and that, in some circumstances, non-cognitive values are epistemically beneficial to theory appraisal. Several questions still need to be addressed, as this final chapter shows. The arguments developed in this dissertation should be seen as a base that can be used to explore these questions.

Bibliography

Ah-King, Malin (2007). "Sexual Selection Revisited – Towards a Gender-Neutral Theory and Practice A Response to Vandermassen's 'Sexual Selection: A Tale of Male Bias and Feminist Denial'", *European Journal of Women's Studies*, 14: 341-348.

Anderson, Elisabeth (2004). "Uses of Value Judgments in Science: A General Argument, with Lessons from a Case Study of Feminist Research on Divorce", *Hypatia*, 19: 1-24.

Antony, Louise 1993. "Quine As Feminist: The Radical Import of Naturalized Epistemology", in Louise Antony and Charlotte Witt (eds.) *A Mind of One's Own: Feminist Essays on Reason and Objectivity* (185-226). Boulder: Westview Press.

Baker, Alan (2003). "Quantitative parsimony and explanatory power", *The British Journal for the Philosophy of Science*, 54(2); 245–259.

Baker, Robin R., and Mark A. Bellis (1993). "Human sperm competition: ejaculate adjustment by males and the function of masturbation", *Animal Behaviour*, 46: 861–885.

Bateman, Angus J. (1948). "Intra-sexual selection in *Drosophila*", *Heredity*, 2(3): 349-368.

Bechtel, William., and Robert C. Richardson (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton, NJ, US: Princeton University Press.

Betzig, Laura (1989). "Causes of conjugal dissolution", *Current Anthropology*, 30: 654–676.

Biology and Gender Study Group (1988). "The Importance of Feminist Critique for Contemporary Cell Biology", *Hypatia*, 3: 61–76.

Brandon, Robert N. (1990). *Adaptation and environment*. Princeton, NJ: Princeton University Press.

Brigandt, Ingo (2015). "Social values influence the adequacy conditions of scientific theories: beyond inductive risk", *Canadian Journal of Philosophy*, 45: 326-356.

Brown, Matthew J. (2013). "Values in Science beyond Underdetermination and Inductive Risk", *Philosophy of Science*, 80: 829-839.

Bryman, Alan (2007). "The research question in social research: What is its role?", *International Journal of Social Research Methodology*, 10(1): 5–20.

Buller, David J. (2005). *Adapting minds*. Cambridge: MIT Press.

Burnham, Kenneth P., and David R. Anderson (2002). *Model Selection and Multimodel Inference*. New York: Springer.

Buss, David M. (1995). "Evolutionary psychology: A new paradigm for psychological science", *Psychological Inquiry*, 6(1): 1–30.

——— (1998). "Sexual strategies theory: Historical origins and current status", *Journal of Sex Research*, 35(1): 19–31.

——— (1999/2008). *Evolutionary Psychology: The New Science of the Mind*. Boston: Allyn and Bacon.

Buss, David M. et al. (1990). "International Preferences in Selecting Mates: A Study of 37 Cultures", *Journal of Cross-cultural Psychology*, 21: 5-47.

Buss, David M., Randy J. Larsen, Drew Westen, and Jennifer Semmelroth (1992). "Sex differences in jealousy: Evolution, physiology, and psychology", *Psychological science*, 3(4): 251-256.

Buss, David M., and David P. Schmitt (1993). "Sexual strategies theory: An evolutionary perspective on human mating", *Psychological Review*, 100: 204–232.

Buss, David M., Martie G. Haselton, Todd K. Shackelford, April L. Bleske, and Jerome C. Wakefield (1998). "Adaptations, exaptations, and spandrels", *American Psychologist*, 53(5): 533–548.

Buss, David M., Todd K. Shackelford, Lee A. Kirkpatrick, and Randy J. Larsen (2001). "A half century of mate preferences: The cultural evolution of values", *Journal of Marriage and Family*, 63(2): 491-503.

Buss, David M., and H. Kern Reeve (2003). "Evolutionary psychology and developmental

dynamics: Comment on Lickliter and Honeycutt (2003)", *Psychological Bulletin*, 129(6): 848–853.

Buss, David M., and David P. Schmitt (2011). "Evolutionary psychology and feminism", *Sex Roles*, 64: 768–787.

Buunk, Braam P., Pietermel Dijkstra et al. (2001). "Age Preferences for mates as related to gender, own age, and involvement level", *Evolution and Human Behavior*, 22: 241-250.

Caporael, Linnda R. (2001). "Evolutionary psychology: Toward a unifying theory and a hybrid science", *Annual review of psychology*, 52(1): 607-628.

Card, David, and Jesse Rothstein (2007). "Racial segregation and the black–white test score gap", *Journal of Public Economics*, 91(11-12): 2158-2184.

Caretti, Vincenzo, Stefano Ciulla and Adriano Schimmenti (2011). "Stalking: Definition of the construct, phenomenological facets, and related behaviors", *Journal of Psychopathology*, 17: 5-12.

Cassidy, Angela (2005). "Popular evolutionary psychology in the UK: an unusual case of science in the media?", *Public Understanding of Science*, 14(2): 115-141.

Coleman, James S. (1966). "Equality of educational opportunity", *Integrated Education*, 6(5): 19-28.

Colombo, Matteo (2017). "Experimental philosophy of explanation rising: The case for a plurality of concepts of explanation", *Cognitive science*, 41(2): 503-517.

Colombo, Matteo, Leandra Bucher, and Yoel Inbar (2016). "Explanatory judgment, moral offense and value-free science", *Review of philosophy and psychology*, 7(4): 743-763.

Colyvan, Mark (1999). "Confirmation theory and indispensability", *Philosophical Studies*, 96(1): 1–19.

Confer, Jaime C., Carin Perilloux, and David M. Buss (2010). "More Than Just a Pretty Face: Men's Priority shifts Toward Bodily Attractiveness in Short-Term Mating Contexts", *Evolution and Human Behavior*, 31: 349-353.

Cosmides, Leda, and John Tooby (1994). "Beyond intuition and instinct blindness: Toward

an evolutionarily rigorous cognitive science”, *Cognition*, 50: 41–77.

Cupach, William R., and Brian H. Spitzberg (1998). “Obsessive relational intrusion and stalking”, in Brian H. Spitzberg and William R. Cupach (eds.), *The dark side of close relationships* (233-263). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

——— (2004). *The dark side of relationship pursuit: From attraction to obsession and stalking*. Mahwah: Lawrence Erlbaum Associates.

Darwin, Charles R. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.

Davis, Keith E., Suzanne C. Swan, and Laura J. Gambone (2012). “Why doesn’t he just leave me alone? Persistent pursuit: A critical review of theories and evidence”, *Sex Roles*, 66: 328-339.

De Smet, Olivia, Tom Loeys, and Ann Buysse (2013). “Ex-couples’ unwanted pursuit behavior: An actor–partner interdependence model approach”, *Journal of Family Psychology*, 27(2): 221-231.

Dorato, Mauro (2004). “Epistemic and nonepistemic values in science”. in Peter Machamer and Gereon Wolters (eds.) *Science, values, and objectivity* (52-77). Pittsburgh: University of Pittsburgh Press.

Douglas, Heather (2000). “Inductive Risk and Values in Science”, *Philosophy of Science*, 67: 559-579.

——— (2005). “Inserting the Public into Science”, in Sabine Maasen and Peter Weingart (eds.) *Democratization of Expertise? Exploring Novel Forms of Scientific Advice in Political Decision-Making, Sociology of the Sciences*, vol. 24 (153-169). Springer Science & Business Media.

——— (2009). *Science, Policy and the Value-Free Ideal*. Pittsburgh: Pittsburgh University Press.

——— (2013). “The value of cognitive values”, *Philosophy of Science*, 80: 796-806.

Downes, Stephen M.(2018). “Evolutionary Psychology”, in Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), URL =

<<https://plato.stanford.edu/archives/fall2018/entries/evolutionary-psychology/>>.

Driver, Julia (2006). *Ethics: The Fundamentals*. Oxford, UK: Blackwell Publishing.

Duntley, Joshua, and David M. Buss (1998). "Evolved anti-homicide modules", Paper presented at the annual meeting of the Human Behavior and Evolution Society, Davis, CA.

——— (2012). "The evolution of stalking", *Sex roles*, 66: 311-327.

Durrant, Russil, and Brian D. Haig (2001). "How to pursue the adaptationist program in psychology", *Philosophical Psychology*, 14(4): 357–380.

Elliott, Kevin C., and Daniel J. McKaughan (2009). "How values in scientific discovery and pursuit alter theory appraisal", *Philosophy of Science*, 76(5): 598-611.

Ellis, Bruce J., and Donald Symons (1990). "Sex differences in sexual fantasy: An evolutionary psychological approach", *Journal of Sex Research*, 27(4): 527-555.

Fausto-Sterling, Anne (1993). "The five sexes: Why male and female are not enough", *The Sciences*, 33: 21-25.

Fedigan, Linda M. (2001). "The Paradox of Feminist Primatology: The Goddess's Discipline?". In Angela Creager, Elizabeth Lunbeck, and Londa Schiebinger (eds.), *Feminism in Twentieth-Century Science, Technology, and Medicine* (46-72). Chicago: University of Chicago Press.

Fehr, Carla (2012). "Feminist Engagement with Evolutionary Psychology", *Hypatia*, 27: 50–72.

——— (2018). "Feminist Philosophy of Biology", in Edward Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition). <https://plato.stanford.edu/archives/fall2018/entries/feminist-philosophy-biology/>

Feyerabend Paul (1975). *Against Method*. London: New Left Books.

——— (1978). *Science in a free society*. London: NLB.

Figueiredo Nascimento, S. Cuccillato, E., Schade, S., & Guimarães Pereira, A. (2016). "Citizen Engagement in Science and Policy-Making". *Reflections and recommendations*

across the European Commission. Luxembourg: Publications Office of the European Union.

Forster, Malcolm, and Elliott Sober (1994). "How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions", *The British Journal for the Philosophy of Science*, 45(1): 1–35.

Frederick, David A., H. Kate St John, Justin R. Garcia, and Elisabeth Lloyd (2018). "Differences in orgasm frequency among gay, lesbian, bisexual, and heterosexual men and women in a US national sample", *Archives of sexual behavior*, 47(1): 273-288.

Fryer Jr, Roland G., and Steven D. Levitt (2004). "Understanding the black-white test score gap in the first two years of school", *Review of Economics and Statistics*, 86(2): 447-464.

Fuller, Steve (2007). *Science v. Religion? Intelligent Design and the Problem of Evolution*. Cambridge: Polity Press.

Funk, Cary, Lee Rainie, and Dana Page (2015). "Public and scientists' views on science and society", *Pew Research Center*.

Gangestad, Steven W., and Randy Thornhill (1997). "The evolutionary psychology of extrapair sex: The role of fluctuating asymmetry", *Evolution and Human Behavior*, 18: 69–88.

Godfrey-Smith, Peter (2001). "Three kinds of Adaptationism", in Steven H. Orzack and Elliott Sober (eds.) *Adaptationism and Optimality* (335–357). Cambridge: Cambridge University Press.

Goldenberg, Maya J. (2015). "Whose Social Values? Evaluating Canada's 'Death of Evidence' Controversy", *Canadian Journal of Philosophy*, 45: 404-424.

Gottschall, Jonathan, Rachel Berkey et al. (2003). "Patterns of characterization in folktales across geographic regions and levels of cultural complexity: Literature as a neglected source of quantitative data", *Human nature*, 14: 365-382.

Gould, Stephen Jay (1978). "Sociobiology: The art of storytelling", *New Scientist*: 530–533.

——— (1997). "Evolution: The pleasures of pluralism", *The New York Review of Books*, June 26.

Gould, Stephen Jay, and Elisabeth S. Vrba (1982). "Exaptation. A missing term in the science of form", *Paleobiology*, 8(1): 4-15.

Gowaty, Patricia Adair (1997). "Sexual dialectics, sexual selection, and variation in reproductive behavior", in *Feminism and evolutionary biology: Boundaries, intersections, and frontiers* (351-384). New York: Chapman and Hall.

Gowaty, Patricia Adair, Yong-Kyu Kim, and Wyatt W. Anderson (2012). "No evidence of sexual selection in a repetition of Bateman's classic study of *Drosophila melanogaster*", *Proceedings of the National Academy of Sciences*, 109(29): 11740-11745.

Green, Sarah (2014). "A philosophical evaluation of adaptationism as a heuristic strategy", *Acta Biotheoretica*, 62: 479-498.

Greenlees, Ian A., and William C. McGrew (1994). "Sex and age differences in preferences and tactics of mate attraction: Analysis of published advertisements", *Ethology and Sociobiology*, 15(2): 59-72.

Griffiths, Paul (1996). "The historical turn in the study of adaptation", *The British Journal for the Philosophy of Science*, 47(4): 511-532.

——— (2001). "From adaptive heuristic to phylogenetic perspective: Some lessons from the evolutionary psychology of emotion", in Harmon R. Holcomb (ed.) *Conceptual challenges in evolutionary psychology* (309-325). Dordrecht: Springer.

Haack, Susan (1993). "Epistemological reflections of an old feminist", *Reason Papers*, 18: 31-43.

Harman, Gilbert (2000). *Explaining Value and other Essays in Moral Philosophy*. Oxford: Clarendon Press.

Haselton, Martie G., and David M. Buss (2000). "Error management theory: A new perspective on biases in cross-sex mind reading", *Journal of personality and social psychology*, 78(1): 81.

——— (2001). "The affective shift hypothesis: The functions of emotional changes following sexual intercourse", *Personal relationships*, 8: 357-369.

——— (2003) "Biases in social judgment: Design flaws or design features?", in Joseph P.

Forgas, Kipling D. Williams, and William von Hippel (eds.), *Social judgments: Implicit and explicit processes* (23-43). New York, NY, US: Cambridge University Press.

Hayden, Michael, Michael Pignone, Christopher Phillips, and Cynthia Mulrow (2002). "Aspirin for the primary prevention of cardiovascular events: a summary of the evidence for the US Preventive Services Task Force", *Annals of Internal Medicine*, 136(2): 161-172.

Hempel, Carl G. (1965). *Aspects of Scientific Explanation*. New York: The Free Press.

Hicks, Daniel J. (2014). "A New Direction for Science and Values", *Synthese*, 191: 3271-3295.

Hintikka, Jaakko (1981). "On the logic of an interrogative model of scientific inquiry", *Synthese*, 47: 69–83.

Holmes, Janet, and Stephanie Schnurr (2005). "Politeness, humor and gender in the workplace: Negotiating norms and identifying contestation", *Journal of Politeness Research. Language, Behaviour, Culture*, 1(1): 121-149.

Hrdy, Sarah (1976). "The Care and Exploitation of Nonhuman Primates by Conspecifics Other than the Mother", *Advances in the Study of Behavior*, VI: 101–158.

——— (1981/1999). *The woman that never evolved*, Cambridge/MA: Harvard University Press.

——— (1986). "Empathy, polyandry, and the myth of the coy female", in Ruth Bleier (ed.) *Feminist approaches to science* (119-146). New York: Teachers College Press.

Hubbard, Ruth. (1990). *The politics of women's biology*. Piscataway/NJ: Rutgers University Press.

——— (1992). "The Political Nature of Human Nature", in Deborah L. Rhode (ed.) *Theoretical Perspectives on Sexual Differences* (140-149). New Haven/CT: Yale University Press.

Intemann, Kristen, and Inmaculada de Melo-Martín (2010). "Social values and scientific evidence: the case of the HPV vaccines", *Biology & philosophy*, 25(2): 203-213.

ISIS-2 (Second International Study of Infarct Survival) Collaborative Group (1988).

"Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2", *Lancet* 2, 8607: 349-60.

Ivani, Silvia (2019). "What We (Should) Talk about When We Talk about Fruitfulness", *European Journal for Philosophy of Science*, 9(1): 4.

Ivanova, Milena (2017). "Aesthetic values in science", *Philosophy Compass* 12(10): e12433.

Jackson, Margaret (1984). "Sex research and the construction of sexuality: A tool of male supremacy?", *Women's Studies International Forum*, 7(1): 43-51.

Jagessar, Jillian DH, and Lorraine P. Sheridan (2004). "Stalking perceptions and experiences across two cultures", *Criminal justice and behavior*, 31(1): 97-119.

Jeffrey, Richard C. (1956). "Valuation and acceptance of scientific hypotheses", *Philosophy of Science*, 23(3): 237-246.

Jensen, Jakob D., Manusheela Pokharel, et al. (2017). "Communicating uncertain science to the public: How amount and source of uncertainty impact fatalism, backlash, and overload", *Risk Analysis*, 37(1): 40-51.

Johnson, Branden B., and Paul Slovic (1995). "Presenting uncertainty in health risk assessment: initial studies of its effects on risk perception and trust", *Risk analysis*, 15(4): 485-494.

——— (1998). "Lay views on uncertainty in environmental health risk assessment", *Journal of Risk Research*, 1(4): 261-279.

Kahan, Daniel M., Donald Braman, Geoffrey L. Cohen, John Gastil, and Paul Slovic (2010). "Who fears the HPV vaccine, who doesn't, and why? An experimental study of the mechanisms of cultural cognition", *Law and human behavior*, 34(6): 501-516.

Kahneman, Daniel, and Amos Tversky (1979). "Prospect theory: An analysis of decision under risk", *Econometrica*, 47(2): 263-292.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler (1991). "The endowment effect, loss aversion, and status quo bias", *Journal of economic perspectives*, 5(1): 193-206.

Ketelaar, Timothy, and Bruce J. Ellis (2000). "Are evolutionary explanations unfalsifiable?

Evolutionary psychology and the Lakatosian philosophy of science", *Psychological Inquiry*, 11(1): 1-21.

Kipling, Rudyard (1902/1989). *Just so stories for little children*. OUP Oxford.

Kitcher, Philip (1989). "Explanatory Unification and the Causal Structure of the World", In Philip Kitcher and Wesley Salmon (eds.) *Scientific Explanation*, (410–505). Minneapolis: University of Minnesota Press.

——— (2003). *Science, truth, and democracy*, Oxford: Oxford University Press.

Knobe, Joshua (2003). "Intentional action and side effects in ordinary language", *Analysis*, 63(3): 190-194.

Koehler, Jonathan J. (1993). "The influence of prior beliefs on scientific judgments of evidence quality", *Organizational behavior and human decision processes*, 56(1): 28-55.

Kourany, Janet A. (2010). *Philosophy of Science after Feminism*. New York: Oxford University Press.

Kuhle, Barry X. (2012). "Evolutionary Psychology is Compatible with Equity Feminism, but Not with Gender Feminism: A Reply to Eagly and Wood (2011)", *Evolutionary Psychology*, 10: 39-43.

Kuhn, Thomas (1962). *The structure of scientific revolution*, Chicago: The University of Chicago.

——— (1977). "Objectivity, value judgement, and theory choice", In *The essential tension. Selected studies in scientific tradition and change* (320–339). Chicago: University of Chicago Press.

——— (1993). "Afterwords", in Paul Horwich (ed.) *World Changes. Thomas Kuhn and the Nature of Science* (311-341). Cambridge MA: MIT Press.

Kukla, Rebecca (2016). "Equipoise, uncertainty, and inductive risk in research involving pregnant women", in Françoise Baylis and Angela Ballantyne (Eds.) *Clinical Research Involving Pregnant Women*. Research Ethics Forum, vol 3 (179-196). Springer, Cham.

Kunda, Ziva (1987). "Motivation and inference: Self-serving generation and evaluation of

evidence", *Journal of Personality and Social Psychology*, 53: 636-647.

——— (1990) "The case for motivated reasoning", *Psychological bulletin* 108(3): 480-498.

Labots, Geert, et al. (2018). "Gender differences in clinical registration trials: is there a real problem?", *British journal of clinical pharmacology*, 84(4): 700-707.

Lacey, Hugh (1999). *Is science value free? Values and scientific understanding*. London: Routledge.

Lakatos, I. (1970). "Falsificationism and methodology of research programmes", in Imre Lakatos and Alan Musgrave (Eds.) *Criticism and the growth of knowledge* (91–195). Cambridge: Cambridge University Press.

Laudan, L. (1982). "Commentary: Science at the Bar - Causes for Concern", *Science, Technology, and Human Values*, 7: 16-19.

——— (2004). "The epistemic, the cognitive, and the social", in Paul Machamer, Geroen Wolters (Eds.), *Science, values, and objectivity* (14–23). Pittsburgh: University of Pittsburgh Press.

Lenard, Philipp (1936). *Deutsche Physik*, Vol. I. München: J.F. Lehmann.

Levi, Isaac (1960). "Must the scientist make value judgments?", *The Journal of Philosophy*, 57(11): 345-357.

Lewens, Tim (2009). "Seven types of adaptationism", *Biology and Philosophy*, 24(2): 161–182.

Lewis, David M., Laith Al-Shawaf, Daniel Conroy-Beam, Kelly Asao, & David M. Buss (2017). "Evolutionary psychology: A how-to guide", *American Psychologist*, 72(4): 353–373.

Lewontin, Richard C. (1998). "The evolution of cognition: Questions we will never answer", in Don Scarborough and Saul Sternberg (Eds.), *An invitation to cognitive science. Methods, models, and conceptual issues: An invitation to cognitive science*, Vol. 4 (106-132). Cambridge, MA, US: The MIT Press.

Lickliter, Robert, and Hunter Honeycutt (2003). "Developmental dynamics: toward a

biologically plausible evolutionary psychology”, *Psychological bulletin*, 129(6): 819.

Lippman, Abby 2006. *The inclusion of women in clinical trials: Are we asking the right questions?*. Toronto, Canada: Women and Health Protection.

Lippman, Julia R. (2018). "I did it because I never stopped loving you: The effects of media portrayals of persistent pursuit on beliefs about stalking", *Communication Research* 45(3): 394-421.

Liu, Brittany S., and Peter H. Ditto (2013). "What dilemma? Moral evaluation shapes factual belief", *Social Psychological and Personality Science*, 4(3): 316-323.

Liu, Katherine A., and Natalie A. Dipietro Mager (2016). "Women's involvement in clinical trials: historical perspective and future implications", *Pharmacy Practice (Granada)* 14.1: 1-9.

Lloyd, Elizabeth A. (2001). "All about Eve: Bias in evolutionary explanations of women's sexuality", in *Proceedings of the Pittsburgh Workshop in History and Philosophy of Biology, Pittsburgh*, 22-46.

——— 2003. "Violence against science: Rape and evolution", in Cheryl Travis (ed.) *Evolution, gender, and rape*. Cambridge, Mass.: MIT Press.

——— (2005). *The case of the female orgasm: Bias in the science of evolution*. Cambridge: Harvard University Press.

——— (2015). "Adaptationism and the logic of research questions: How to think clearly about evolutionary causes", *Biological Theory*, 10(4): 343–362.

Longino, Helen. (1990). *Science as Social Knowledge. Values and Objectivity in Scientific Inquiry*. Princeton/NJ: Princeton University Press.

——— (1995). "Gender, politics, and the theoretical virtues", *Synthese*, 104(3): 383–397.

——— 2002. *The Fate of Knowledge*. Princeton, N.J., and Oxford: Princeton University Press.

——— (2008). "Values, heuristics, and the politics of knowledge", Ii Carrier, Martin, Don Howard, and Janet Kourany (eds.) *The Challenge of the Social and the Pressure of Practice*:

Science and Values Revisited (68-86). Pittsburgh, PA: University of Pittsburgh Press.

Lord, Charles G., Lee Ross, and Mark R. Lepper (1979). "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence", *Journal of personality and social psychology*, 37(11): 2098-2109.

Lundstedt, Sven (1968). "A note on asking questions", *The Journal of General Psychology*, 79(2): 229–239.

Lyndon, Amy E., H. Colleen Sinclair, John MacArthur, Ben Fay, Ellen Ratajack, and Katie Collier (2012). "An introduction to issues of gender in stalking research", *Sex Roles*, 66(5-6): 299-310.

Machery, Edouard (2011). "Discovery and confirmation in evolutionary psychology", in Jesse J. Prinz (Ed.), *The Oxford handbook of philosophy of psychology* (232–246). Oxford: Oxford: Oxford University press.

Machery, Edouard, and H. Clark Barrett (2006). "David J. Buller: Adapting Minds: Evolutionary Psychology and the Persistent Quest for Human Nature essay review: Debunking adapting minds", *Philosophy of Science*, 73(2): 232–246.

Machery, Edouard, Stephen Stich, David Rose, Amita Chatterjee, Kaori Karasawa, Noel Struchiner, Smita Sirker, Naoki Usui, and Takaaki Hashimoto (2015). "Gettier Across Cultures", *Noûs* 51(3): 645-664.

Mason, Jennifer (2002). *Qualitative researching*. London: Sage Publications.

McMullin, Ernan (1976). "The fertility of theory and the unit for appraisal in science", in Robert S. Cohen, Paul K. Feyerabend, and Marx W. Wartofsky (eds.), *Essays in memory of Imre Lakatos* (395–432). Dordrecht: Springer Netherlands.

——— (1979). "The ambiguity of historicism", in Peter D. Asquith and Henry Ely Kyburg (eds.), *Current research in philosophy of science* (55–83). Lansing: Philosophy of Science Association.

——— (1983). "Values in science", *PSA: Proceedings of the biennial meeting of the philosophy of science association*, 1982(2): 3–28.

——— (1985). "Galilean idealization", *Studies in History and Philosophy of Science Part A*,

16(3), 247–273.

——— (2009). “The virtues of a good theory”, in Stathis Psillos and Martin Curd (eds.), *The Routledge companion to philosophy of science* (498–508). London: Routledge.

McPhillips, Kathryn, Virginia Braun, and Nicola Gavey (2001). “Defining (hetero) sex: how imperative is the “coital imperative”?”, *Women's Studies International Forum*, 24(2): 229–240.

Meloy, John Reid. (1999). “Stalking: An old behavior, a new crime”, *Psychiatric Clinics of North America*, 22(1): 85–99.

Meloy, J. Reid, and Helen Fisher (2005). “Some thoughts on the neurobiology of stalking”, *Journal of Forensic Science*, 50(6): 1472–1480.

Morrison, Margaret (2000). *Unifying scientific theories: Physical concepts and mathematical structures*. Cambridge University Press.

Munro, Geoffrey D., and Peter H. Ditto (1997). “Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information”, *Personality and Social Psychology Bulletin*, 23(6): 636–653.

Nature Editorial (2004). “Going Public”, *Nature*, 431: 833.

Nesse, Randolph M. (1990). “Evolutionary explanations of emotions”, *Human Nature*, 1: 261–289.

Nesse, Randolph M., and Stephen C. Stearns (2008). “The great opportunity: Evolutionary applications to medicine and public health”, *Evolutionary Applications*, 1(1): 28–48.

Nickerson, Raymond S. (2000). “Null hypothesis significance testing: a review of an old and continuing controversy”, *Psychological methods*, 5(2): 241–301.

Nolan, Daniel (1997). “Quantitative parsimony”, *The British Journal for the Philosophy of Science*, 48(3): 329–343.

——— (1999). “Is fertility virtuous in its own right?”, *The British Journal for the Philosophy of Science*, 50(2): 265–282.

Norton, John D. (Unpublished). "Epistemic Virtues and Epistemic Values: A Skeptical Critique", in *The Material Theory of Induction*. https://www.pitt.edu/~jdnorton/papers/material_theory/5.%20Epistemic_Virtues.pdf

O'Neill, Elizabeth, & Edouard Machery (2014). "Experimental Philosophy", in *Current controversies in experimental philosophy* (vii-xxix). New York: Routledge.

Öhman, Arne, Anders Flykt, and Francisco Esteves (2001). "Emotion drives attention: detecting the snake in the grass", *Journal of experimental psychology*, 130(3): 466-478.

Pedersen, William C. et al. (2011). "Are Men and Women Really That Different? Examining Some of Sexual Strategies Theory (SST)'s Key Assumptions about Sex-Distinct Mating Mechanisms", *Sex Roles*, 64: 629–643.

Peirce, Charles S. (1935). *The Collected Papers of Charles Sanders Peirce*, volume I-VI. Harvard University Press, Cambridge, Mass.

Perrin, Paul B. et al. (2011). "Aligning Mars and Venus: The Social Construction and Instability of Gender Differences in Romantic Relationships", *Sex Roles*, 64: 613–628.

Pinker, Stephen (2002). *The Blank slate: The modern denial of human nature*. New York: Viking.

Pinn, Vivian W (2003). "Sex and gender factors in medical studies: implications for health and clinical practice", *Jama*, 289.4: 397-400.

Pitt, Joseph C. (2001). "The dilemma of case studies: Toward a Heraclitian philosophy of science", *Perspectives on Science*, 9(4): 373-382.

Planck, Max (1915). *Eight lectures on theoretical physics*. New York: Columbia University Press.

Poincaré, Henri (1905). *Science and hypothesis*. New York: Walter Scott Publishing.

Psillos, Stathis (2015). "Evidence: Wanted, Alive or Dead", *Canadian Journal of Philosophy*, 45: 357-381.

Quine, Willard V. O. (1955). "Posits and reality", in *The Ways of Paradox and Other Essays* (233–241). Cambridge: Harvard University Press.

Reed, Evelyn (1978). *Sexism and science*. New York: Pathfinder Press.

Reddy, Cristopher (2009). "Scientist Citizens", *Science*, 323: 1405. 10.1126/science.1173003

Reiss, Julian and Jan Sprenger (2017). "Scientific Objectivity", in Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), URL = <<https://plato.stanford.edu/archives/win2017/entries/scientific-objectivity/>>.

Resnik, David B. (1991). "How-possibly explanations in biology", *Acta Biotheoretica*, 39(2): 141-149.

Retzbach, Andrea, and Michaela Maier (2015). "Communicating scientific uncertainty: Media effects on public engagement with science", *Communication Research*, 42(3): 429-456.

Richardson, Robert C. (2007). *Evolutionary psychology as maladapted psychology*. Cambridge: MIT Press.

Richardson, Sarah S. (2008). "When gender criticism becomes standard scientific practice: the case of sex determination genetics", in Londa Schiebinger (Ed.) *Gendered Innovations in Science and Engineering*, (22-42). Palo Alto, CA: Stanford University Press.

Ridker, Paul M., Nancy R. Cook et al. (2005). "A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women", *New England Journal of Medicine*, 352(13): 1293-1304.

Rolin, Kristina (1998). "What should a normative theory of values in science accomplish?", *1998 Meeting of the Philosophy of Science Association*, posted at: <http://scistud.umkc.edu/psa98/papers>.

Rooney, Phyllis (1992). "On Values in Science: Is the Epistemic/Non-Epistemic Distinction useful?", *Proceedings of the Biennial Meetings of the PSA 1992*: 13-22.

——— (2017). "The Borderlands Between Epistemic and Non-Epistemic Values", in Kevin C. Elliott and Daniel Steel (eds.) *Current Controversies in Values and Science* (31-45). New York: Taylor and Francis.

Rudner, Richard (1953). "The Scientist Qua Scientist Makes Value Judgments", *Philosophy*

of Science, 20: 1-6.

Ruse, Michael (1999). *Mystery of mysteries: Is evolution a social construction?*. Cambridge, Mass.: Harvard University Press.

Salmon, Wesley C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Samuelson, William, and Richard Zeckhauser (1988). "Status quo bias in decision making", *Journal of risk and uncertainty*, 1(1): 7-59.

Sanchez, Diana T., Jennifer Crocker, and Karlee R. Boike (2005). "Doing gender in the bedroom: Investing in gender norms and the sexual experience", *Personality and Social Psychology Bulletin*, 31(10): 1445-1455.

Scher, Steven J., and Frederick Rauscher (2003). *Evolutionary psychology: Alternative approaches*. Boston: Kluwer.

Schiebinger, Londa (1999). *Has Feminism Changed Science?* Cambridge: Harvard University Press.

Schindler, Samuel (2014). "A matter of Kuhnian theory-choice? The GWS model and the neutral current", *Perspectives on Science*, 22(4): 491-522.

——— (2017). "Theoretical fertility McMullin-style", *European Journal for Philosophy of Science*, 7(1): 1–23.

Schmitt, David P., and David M. Buss (2001). "Human mate poaching: Tactics and temptations for infiltrating existing mateships", *Journal of personality and Social Psychology*, 80(6): 894-917.

Schmitt, David P. et al. (2003). "Universal Sex Differences in the Desire for Sexual Variety: Tests from 52 Nations, 6 Continents, and 13 Islands", *Journal of Personality and Social Psychology*, 85: 85-104.

Schmitt, David. P., & June J. Pilcher (2004). "Evaluating evidence of psychological adaptation: How do we know one when we see one?", *Psychological Science*, 15(10): 643–649.

Schmitt, David P., et al. (2012). "A reexamination of sex differences in sexuality: New studies reveal old truths", *Current Directions in Psychological Science*, 21(2): 135-139.

Schroeder, S. Andrew (2017). "Using Democratic Values in Science: an Objection and (Partial) Response", *Philosophy of Science*, 84(5): 1044-1054.

Segall, Robert (2009). *The fertility of theories*. Doctoral dissertation: University of Cape Town.

Schupbach, Jonah N., and Jan Sprenger (2011). "The logic of explanatory power", *Philosophy of Science*, 78(1): 105-127.

Sheridan, Lorraine, Raphael Gillett & Graham Davies (2002). "Perceptions and prevalence of stalking in a male sample", *Psychology, Crime & Law*, 8(4): 289-310.

Shorey, Ryan C., Tara L. Cornelius, and Catherine Strauss (2015). "Stalking in college student dating relationships: A descriptive investigation", *Journal of family violence*, 30(7): 935-942.

Simon, Viviana (2005). "Wanted: women in clinical trials", *Science*, 10 Jun 2005, 308(5728): 1517-1517.

Singh, Devendra (1993). "Adaptive significance of female physical attractiveness: role of waist-to-hip ratio", *Journal of personality and social psychology*, 65(2): 293-307.

Sjoberg, Laura, and Caron E. Gentry (2007). *Mothers, monsters, whores: women's violence in global politics*. Zed Books.

Slimak, Michael W., and Thomas Dietz (2006). "Personal values, beliefs, and ecological risk perception", *Risk analysis* 26(6): 1689-1705.

Smiler, Andrew P. (2011). "Sexual strategies theory: Built for the short term or the long term?", *Sex Roles*, 64: 603-612.

Smith, Daniel (2013). "Mathematical Myopia: Debunking Common Six Sigma Misconceptions", *Amstatnews* (November 1 2013).
<https://magazine.amstat.org/blog/2013/11/01/mathmyopia/>

Smith, Sharon G., Xinjian Zhang, Kathleen C. Basile, Melissa T. Merrick, Jing Wang,

Marcie-jo Kresnow, and Jieru Chen (2018). "The national intimate partner and sexual violence survey (NISVS): 2015 Data Brief – Updated Release". Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Sober, Elliott (2000). *Philosophy of biology*. Boulder: Westview Press.

Sommers, Christina Hoff (1994). *Who Stole Feminism? How Women Have Betrayed Women*. New York: Simon & Schuster.

Spitzberg, Brian H., and William R. Cupach (2007). "The state of the art of stalking: Taking stock of the emerging literature", *Aggression and violent Behavior*, 12(1): 64-86.

Staley, Kent W. (2017). "Decisions, Decisions", in Kevin C. Elliott and Ted Richards (eds.) *Exploring Inductive Risk: Case Studies of Values in Science* (37-55). New York: Oxford University Press.

Steel, Daniel (2010). "Epistemic values and the argument from inductive risk", *Philosophy of Science*, 77(1): 14–34.

Steel, Daniel, Chad Gonnerman, and Michael O'Rourke (2017). "Scientists' attitudes on science and values: Case studies and survey methods in philosophy of science", *Studies in History and Philosophy of Science Part A*, 63: 22-30.

Steering Committee of the Physicians' Health Study Research Group. (1989). "Final report on the aspirin component of the ongoing Physicians' Health Study", *New England Journal of Medicine*, 321(3): 129-135.

Sterelny, Kim, and Paul Griffiths (1999). *Sex and death*. Chicago: The University of Chicago Press.

Surbey, Michele K., and Colette D. Conohan (2000). "Willingness to engage in casual sex", *Human Nature*, 11(4): 367-386.

Symons, Donald (1979). *The Evolution of Human Sexuality*. New York: Oxford University Press.

Tate, Chuck (2011). "The 'Problem of Number' Revisited: The Relative Contributions of Psychosocial, Experiential, and Evolutionary Factors to the Desired Number of Sexual Partners", *Sex Roles*, 64: 644-657.

Thornhill, Randy, Stevem W. Gangestad, and Randall Comer (1995). "Human female orgasm and mate fluctuating asymmetry", *Animal Behaviour*, 50: 1601–1615.

Tjaden, Patricia, and Nancy Thoennes (1998). "Prevalence, Incidence, and Consequences of Violence against Women: Findings from the National Violence against Women Survey. Research in Brief". Department of Justice, Washington, DC. National Inst. of Justice.; Centers for Disease Control and Prevention (DHHS/PHS), Atlanta, GA.

Tooby, John, and Leda Cosmides (1990). "The past explains the present: Emotional adaptations and the structure of ancestral environments", *Ethology and Sociobiology*, 11: 375–424.

——— (1992). "The psychological foundations of culture". In J. Barkow, L. Cosmides, J. Tooby (eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (19–136). New York: Oxford University Press.

——— (2005), "Conceptual Foundations of Evolutionary Psychology". In David Buss (ed.) *The Handbook of Evolutionary Psychology* (5-67). Hoboken, NJ: Wiley.

Trivers, Robert L. (1972). "Parental Investment and Sexual Selection", in Bernard Campbell (ed.) *Sexual Selection and the Descent of Man* (136–179). Chicago: Aldine de Gruyter.

Tulodziecki, Dana (2011). "A case study in explanatory power: John Snow's conclusions about the pathology and transmission of cholera", *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 42(3): 306-316.

Tybur, Joshua M., Debra Lieberman, and Vladas Griskevicius (2009). "Microbes, mating, and morality: individual differences in three functional domains of disgust", *Journal of personality and social psychology*, 97(1): 103-122.

Valles, Sean A. (2012). "Heterogeneity of risk within racial groups, a challenge for public health programs", *Preventive medicine*, 55(5): 405-408.

Vandermassen, Griet (2004). "Sexual Selection. A tale of Male Bias and Feminist Denial", *European Journal of Women's Studies*, 11: 9-26.

——— (2005). *Who's afraid of Charles Darwin? Debating feminism and evolutionary theory*. Lanham: Rowan & Littlefield.

Van Fraassen, Bas C. (1980). *The scientific image*. Oxford: Oxford University Press.

Walker, Caren M., Tania Lombrozo, Christine H. Legare, and Alison Gopnik (2014). "Explaining prompts children to privilege inductively rich properties", *Cognition*, 133: 343–357.

Wasserman, Larry (2013). *All of statistics: a concise course in statistical inference*. Springer Science & Business Media.

Williams, Joseph J., and Tania Lombrozo (2010). "The role of explanation in discovery and generalisation: Evidence from category learning", *Cognitive Science*, 34: 776–806.

Wimsatt, William C. (2007). *Re-engineering philosophy for limited beings*. Cambridge: Harvard University Press.

Woodward, James (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Yanowitz, Karen L. (2006). "Influence of gender and experience on college students' stalking schemas", *Violence and Victims*, 21(1): 91-100.

Zahar, Elie. (1973). "Why did Einstein's Programme supersede Lorentz's? (II)", *The British Journal for the Philosophy of Science*, 24(3): 223-262.

Acknowledgments

My sincere gratitude goes to my supervisors Prof. Jan Sprenger and Dr. Matteo Colombo. Working with Jan and Matteo has been a great pleasure and an exciting adventure. I have learned a lot. I wish to thank them for their invaluable advice, support, and encouragement.

I am honoured to have Prof. Catarina Dutilh Novaes, Prof. Maureen Sie, Prof. Henk de Regt, and Dr. Leah Henderson as members of the committee. I am very grateful to all of them and I feel very lucky to have leading researchers of my field on the committee.

Special thanks go to my *paranimfen*. Colin and I shared two offices and a lot of food. We were colleagues but most of all we were true friends (and we still are!). Colin, you are a constant source of inspiration, thank you! Maria, during our master studies in Genoa we dreamt of being PhD students, while reading about rigid designators, inference to the best explanation, and Gödel. What's next? Thanks for your support and friendship. We will meet in Piazza Lavagna, as usual.

Thanks to my colleagues: Alessandra, Alfred, Willem, Hanno, Lasha, Machteld, Chiara, Georgie, Bart, Nathan, Dominik, Barbara, and Jun. I have enjoyed every seminar, conference, lunch, drink, chat, coffee, party we had. I have been very fortunate to meet you. Special thanks to Amanda Cawston, who kindly accepted the role of junior research supervisor. Her expertise and kindness have been crucial in the final year of my PhD.

Some chapters of this dissertation were produced in collaboration with others. I am very grateful to my co-authors Jan, Matteo, and Dr. Leandra Bucher. And I wish to thank Prof. Nicla Vassallo, Prof. Marcello Frixione, and Dr. Maria Cristina Amoretti for their support and encouragement from Genoa.

I am grateful to the Philosophy Department of Tilburg University for supporting my participation in conferences and visit to the University of Sydney. I wish to thank the Theory and Method in Biosciences group at the University of Sydney and Prof. Paul Griffiths for making my stay such an exciting experience.

I am thankful to my family: Marco, Nadia, Laura, Lucia, Arturo. They cannot be in Tilburg today, but they are always with me. Thanks for your support and encouragement.

I am surrounded by the most beautiful, kind, fun friends and I am extremely grateful to them: Rossana, Noemi, Luca, Mattia, Elisa, Davide, Maria, Tania, Atnan, Anna, Lucrezia, Astrid,

Lucia, Carlo, Joanne, Denise, Cesare, Laura, Valentino, and Tessa. They taught me so much.
Thank you!

