Biases and fallacies: The role of motivated irrationality in fallacious reasoning

Vasco Correia
Institute for the Philosophy of Language
New University of Lisbon

Abstract: This paper focuses on the effects of motivated biases on the way people reason and debate in everyday life. Unlike heuristics and cognitive (or 'cold') biases, motivated (or 'hot') biases are typically caused by the influence of a desire or an emotion on the cognitive processes involved in judgmental and inferential reasoning. In line with the ‘motivational’ account of irrationality, I argue that these biases are the cause of a number of fallacies that ordinary arguers commit unintentionally, particularly when the commitment to a given viewpoint is very strong. Drawing on recent work in argumentation theory and psychology, I show that there are privileged links between specific types of biases and specific types of fallacies. This analysis provides further support to the idea that people’s tendency to arrive at the desired conclusions hinges on their ability to construct plausible justifications for those conclusions. I suggest that this effort to rationalize unintentionally biased views is the reason why the correlative fallacies tend to be so persuasive.

Keywords: argumentation, biases, confirmation bias, fallacies, focusing illusion, hasty generalization, one-sidedness, straw man

1. Introduction

According to Hamblin’s influential definition, “a fallacious argument … is one that seems to be valid but is not so” (Hamblin 1970: 12). Even though some aspects of this conception have been questioned in recent research (Johnston 1995, Eemeren et al. 2009), it has the merit of highlighting the treacherous nature of fallacies. Yet, fallacies can be deceitful in two very different senses: in the sense that they can be committed intentionally to deceive an audience, to be sure, but also in the sense that they can be committed unintentionally by honest arguers. To that extent, as Tindale (2007: 15) suggests, “we should not take deception to be part of the definition of fallacy”. Despite the common association between fallacies and treachery, ordinary arguers often commit fallacies accidentally, either because they are mislead by the resemblance between correct and incorrect arguments with a similar structure, or because their attachment to a given position is so strong that it becomes tempting to ‘jump to conclusions’. As Walton observes, “[m]any fallacies are committed because the proponent has such strong interests at stake in putting forward a particular argument, or is so fanatically committed
to the position advocated by the argument, that she is blind to weaknesses in it that would be apparent to others not so committed. In this kind of case the deception may not quite be intentional, because the proponent doesn’t see the argument as faulty” (Walton, in press).

In fact, the profusion of empirical studies conducted by psychologists from the 1960s onwards reveals that unintentional fallacies may not always be as ‘accidental’ as one might assume. Granted, there are random lapses that lead to fallacies in everyday life debates: the arguer may be tired, inexperienced, lack concentration, and so forth. However, experiments have consistently shown that most people tend to commit specific errors of reasoning which appear to be too systematic and too pervasive to be random (heuristics and biases). This paper focuses specifically on motivated biases and their relationship with informal fallacies. While heuristics are ‘mental shortcuts’ that provide intuitive solutions that are “in general quite useful, but sometimes lead to severe and systematic errors” (Tverky and Kahneman 1974: 1124), biases are systematic errors that invariably distort the subject’s reasoning and judgment. More importantly, heuristics are generally ‘cold’ (or non-motivated) causes of irrationality, whereas biases are often ‘hot’ (or motivated), i.e., induced by a particular desire or emotion. Bearing in mind that emotions often come into play in real life debates, it seems paramount to examine to what extent, and in which ways, motivated biases can lead us to commit fallacies unintentionally.

The present article purports to elucidate some of the known effects of motivated biases on argumentation by combining two different approaches: on the one hand, the critical analysis of informal fallacies proposed by argumentation theorists, and, on the other hand, the empirical research conducted by psychologists on irrational inference. In section 2, I analyze one of the most common biases in argumentation – the ‘confirmation bias’ – and argue that it induces irrational reasoning through its impact both on the process of evidence seeking (selective evidence gathering) and on the choice of premises (selective representation of relevance). I also show that these effects account for the empirical evidence demonstrating that the confirmation bias reinforces the phenomenon of polarization of opinions. In section 3, I put forward the hypothesis that the ‘focusing illusion’ is one of the main sources of hasty generalization and one-sidedness, particularly since it leads the arguers to overlook some of the available evidence. Furthermore, I suggest that the focusing illusion also gives rise to unintentional cases of the Straw Man fallacy, in which the arguer interprets the opponent’s view in an unconsciously biased fashion. In section 4, I explore the connections between some paradigmatic cases of motivated irrationality (wishful thinking, self-deception and rationalization) and three specific fallacies (argument from consequences, slothful induction and misidentification of the cause).

2. Selective choice of premises, one-sidedness and the confirmation bias

The question of the causes of fallacious reasoning has been the source of much controversy over the past decades. The proponents of the ‘cognitive’ account maintain

---

1 For a fuller analysis of the distinction between ‘hot cognition’ and ‘cold cognition’ see Tetlock & Levi (1980) and Kunda (1990).
that inferential errors stem essentially from cognitive malfunctioning (Nisbett & Ross 1980), whereas the proponents of the ‘motivational’ account maintain that the basic cause of irrational reasoning is an emotion or a motive, and more exactly the impact that affective states seem to have on cognitive processes (Kunda 1990, Baron 1988, Gilovich 1991). Yet, as Tetlock and Levi (1980: 83) suggest, “neither side is likely to ‘win’ the cognition-motivation debate”, and “a more profitable strategy is to focus on identifying and clarifying conceptual ambiguities in the cognitive and motivational positions”. From that perspective, the challenge is not to determine whether motivational biases are more or less predominant than cognitive biases, but to examine how exactly those biases operate, and what specific effects they have on the structure of our reasonings.

In the context of argumentation, the phenomenon of motivated irrationality tends to occur when the arguer ‘feels very strongly’ about a given standpoint: the arguer’s commitment to that standpoint, via the goals and interests it solicits, is susceptible to bias the way she thinks and argues on the issue. That is at least what an impartial listener might be able to notice, since the arguer himself is usually unaware of her own partiality. In that respect, it is important to stress that the effects of unconscious biases on argumentation are very similar to those observed in deliberately biased argumentation, affecting both the process of evidence gathering (partiality) and the very interpretation of the selected evidence (distortion). This seems to explain, in particular, why it is so difficult to know whether a given argument is intentionally biased or not, given that the effects of biases on the discourse are identical in both cases. Hence, for example, if a lawyer produces a biased account of her client’s actions in front of the court, it is virtually impossible to determine whether that lawyer is deliberately trying to manipulate the audience, or if the desire to win the case (or the sympathy toward the client, etc.) unconsciously affects her judgment, thereby causing her to believe sincerely in the client’s version of events. All we know with a certain degree of certainty, thanks to the work of psychologists, is that people’s judgments and inferences are frequently biased both unconsciously and involuntarily (Pohl 2004: 2).

One of the main effects of cognitive biases on argumentation is the phenomenon of selective choice of premises. Even without realizing it, arguers often tend to put forward a set of premises that provide uniquely reasons in support of the claim they which to make. While this sort of one-sidedness may not always be fallacious, it seems in many cases to undermine the rationality and the credibility of the argument. Perelman and Olbrechts-Tyteca highlight that problem in The New Rhetoric:

[All argumentation is selective. It chooses the elements and the method of making them present. By doing so it cannot avoid being open to accusations of incompleteness and hence of partiality and tendentiousness. (Perelman and Olbrechts-Tyteca 1969: 119)]

Whenever this phenomenon occurs unintentionally, the selective choice of premises seems to derive from the process of selective evidence gathering that is often linked to the confirmation bias (Kunda 1999: 115, Baron 1988: 282). The confirmation bias, which according to Baron (1988: 280) constitutes “the major bias in thinking about beliefs of all types”, can be defined as the tendency to favor information that confirms one’s own preconceptions. Under the influence of a given desire or emotion, the arguer tends to
focus on the evidence that seems to confirm his claim and, conversely, to overlook the evidence that seems to disconfirm it. In principle, the affect that causes the biased search for evidence is associated to the arguer’s commitment towards the credited position. Oswald and Grosjean (2004: 81) suggest that “this tendency exists … because the possibility of rejecting the hypothesis is linked to anxiety or other negative emotions”. For example, a person who feels very strongly against the use of transgenic wheat may tend to evoke uniquely scientific studies that confirm her claims, and to disregard studies seemingly incongruent with it. But conversely, the motivating affect could also be the desire to be right about the issue at stake, as it seems to happen when people debate a certain issue just for the sake of debating. In such cases, it is the emotion of pride that comes into play, similarly to what happens with self-serving biases, leading the arguer to seek for information supportive of her own position. As a result, the premises put forward might be so partial that the argument loses its balance and credibility, thereby undermining the goal of the discussion. As Perelman and Olbrechts-Tyteca (1969: 65) observe, “[h]is hearers may refuse their adherence [because] they may see that his choice of premises is one-sided, or they may be shocked by the tendentious way in which the premises were advanced”.

An additional danger of the confirmation bias is the phenomenon of ‘attitude polarization’. This aspect was confirmed beyond expectations by a famous experiment conducted by Lord, Ross and Lepper (1979). The researchers exposed subjects supporting and opposing death penalty to two purported studies, one seemingly confirming and one seemingly disconfirming their beliefs regarding the deterrent efficacy of the capital punishment. Predictably, they found that both proponents and opponents of the death penalty rated the studies that confirmed their existing views as the most convincing and probative ones. Less predictably, though, they found that the ‘pro’ subjects became even more favorable to the capital punishment after being exposed to the information, while the ‘anti’ subjects became even more opposed to it. In other words, the polarization of opinions seemed to have increased after exposure to information, despite the fact that the body of evidence was precisely the same:

If our study demonstrates anything, it surely demonstrates that social scientists cannot expect rationality, enlightenment, and consensus about policy to emerge from their attempts to furnish ‘objective’ data about burning social issues. If people of opposing views can each find support for those views in the same body of evidence, it is small wonder that social science research, dealing with complex and emotional social issues and forced to rely upon inconclusive designs, measures, and modes of analysis, will frequently fuel rather than calm the fires of debate. (Lord, Ross and Lepper 1979: 2108)

Thus, the confirmation bias seems to provide a remarkable example of the way cognitive illusions affect our reasoning during a discussion. The point to be made here is that our well-intended efforts to provide good reasons in support of our claims may be compromised by a process of selective evidence seeking and of biased interpretation that we are not even aware of. Oswald and Grosjean (2004: 79) also insist on that point: “Here, the issue is not the use of deceptive strategies to fake data, but forms of interpretation processing that take place more or less unintentionally”. Even if we
genuinely try to reach a balanced argument, based on reliable evidence and on correct
inferences, there is still a significant risk that we tend to focus too much on a partial
aspect of the available information and end up putting forward a tendentious justification
of our view. Moreover, there is evidence that a higher degree of experience or expertise
on a given matter does not suffice to counteract this type of effect. As Koehler et al.
(2002: 692) explain: “Experts might be more or less susceptible to confirrnatory bias than
novices. On the one hand, their extensive experience may help them learn to evaluate
evidence in a more impartial manner, and could also lead them to spontaneously generate
alternatives to the focal hypothesis. On the other hand, the broader knowledge base of
experts might lead them to more readily generate a biased set of reasons favoring the
focal hypothesis”.

Finally, it is worth noting that the confirmation bias is sometimes aggravated by the
well-documented belief bias, which is generally described as a tendency to endorse
arguments whose conclusions seem believable (Evans 2004: 136). In an interesting study,
Evans et al. (1983) were able to demonstrate that people are significantly influenced by
whether or not they agree with the conclusions of an argument. Subjects were presented
with a series of syllogisms and asked to assess their validity. There were in fact four
different categories of syllogisms, which the researchers classified as Valid-Believable,
Valid-Unbelievable, Invalid-Believable and Invalid-Unbelievable. As predicted, the
degree of ‘believability’ of the conclusions seemed to affect the way the subjects assessed
the syllogisms’ validity. In particular, it appeared that the acceptance rate was much
higher for believable syllogisms than for unbelievable ones, both for valid and invalid
cases. And conversely, they tended to reject valid arguments with unbelievable
conclusions, presumably because the unlikelihood of the conclusion biased their
evaluation of the logical strength of the inference. According to Evans, this phenomenon
arises from “a conflict between two types of thought processes, one reasoning logically
according to the instructions and the other prompting people to respond on the basis of
their prior beliefs” (Evans 2004: 139-140)². Even though the belief bias is a ‘cold’
heuristic that characterizes the way our cognitive processes are structured, it seems to
reinforce ‘hot’ occurrences of the confirmation bias, in which the conclusions of the
arguments typically are motivationally supported hypotheses, that is, “those hypothesis
with respect to which there are positive or negative emotions depending on the outcome
of the test” (Oswald and Grosjean 2004: 90). In both cases, individuals tend to reason in a
way that favors the confirmation of their original positions.

3. Focusing illusion, hasty generalization and relevance fallacies

The one-sidedness of an argument has sometimes less to do with the partiality of the
sources than with the blinkered viewpoint adopted by the arguer in the process of
interpreting the information. On many occasions, the problem seems to be that the arguer
focuses too much on one aspect of the issue. This corresponds to what Schkade and
Kahneman call the focusing illusion: “When a judgment about an entire object or
category is made with attention focused on a subset of that category, a focusing illusion is

² For a detailed account of the “dual process theory of reasoning”, see J. Evans, & D.E. Over (1996)
likely to occur, whereby the attended subset is overweighed relative to the unattended subset” (Schkade and Kahneman 1998: 340). If a person looks only at the negative consequences of a given decision, for instance, completely disregarding its possible benefits, she is likely to conclude that it is preferable not to take that decision. Pirie (2006: 121-122) considers an example of this kind:

*I’m not going to get married. There would be all that extra responsibility, not to mention the loss of my freedom. Think of the costs of raising children and putting them through college. Then there are the increased insurance premiums…*

This example is interesting because the subject’s selective focusing seems to be motivated by the anxiety of getting married and having children. Granted, there is nothing irrational about taking that anxiety into account during the process of practical reasoning. But if that anxiety causes the subject to place too much importance on the negative aspects of marriage, the reasoning through which he weighs the pros and the cons of marriage is likely to be tendentious. Walton also stresses this point: “an argument is more plausible if it is based on a consideration of all the evidence in a case, on both sides of the issue, than if it is pushing only for one side and ignoring all the evidence, even if it may be good evidence, on the other side. So if an argument is biased, that is, if it pushes only for one side, we discount that argument as being worthless (Walton 2006: 238).

Furthermore, the focusing illusion seems to play an important role in our tendency to over-generalize and to commit the fallacy of *hasty generalization*, that is, to make general statements that “conclude more than is warranted by the evidence” (Tindale 2007: 151). In particular, if our attention is focused exclusively on one aspect of things, we may plausibly overlook relevant exceptions. Many authors alert to the dangers of this phenomenon. Walton, for example, considers it to be an abundant source of stereotypes, prejudices and black-and-white statements:

A serious problem with argumentation based on generalizations is that some people who are passionately committed to a viewpoint tend to overlook qualifications that are needed in a specific case … Such a lack of flexibility in argumentation and insensitivity to a possible need for qualifying a generalization is at the root of the rigid stereotyping that is characteristic of fanatic and dogmatic arguers who are intensely committed to their convictions (Walton 2006: 20)

Interestingly, one of the experiments designed to demonstrate the focusing illusion is built around a particular stereotype. Schkade and Kahneman (1998) asked Californians and Midwesterners who they thought was happier. Not surprisingly, both Californians and Midwesterners said that Californians must be considerably happier. Yet, the self-reported life satisfaction was roughly the same in both regions, which indicates that the subjects’ assessments were based on an illusion. According to the authors, the bias stems from the fact that people focus on the sunny weather and easy-going life-style of California, and tend to neglect other important aspects of life: “There appears to be a stereotyped perception that people are happier in California. This perception is anchored in the perceived superiority of the California climate … Nevertheless, contrary to the intuitions of our respondents, the advantages of life in California were not reflected in
differences in the self-reported overall life satisfaction of those who live there” (Schkade and Kahneman 1998: 345). The focusing illusion can have dangerous effects on people’s beliefs and decisions, not only because the stereotypes in question are often negative and prejudicial, but also because it may lead to unjustified initiatives: for example, people might actually want to move to California with the illusion that it would make them happier.

Relevance fallacies could also be induced by the focusing illusion, particularly when the arguer’s attention is biased by a vested interest. This hypothesis seems to account for unintentional occurrences of the Straw Man fallacy, which are arguably more frequent than one might assume. The straw man fallacy is generally described as an intentional strategy whereby the arguer misrepresents or distorts the opponent’s position in such a way that it becomes easier to refute, and then refutes that misrepresentation instead of the opponent’s actual position (Talisse and Aikin 2006, Johnson and Blair 1983, Walton 1989). However, there seem to be unintentional cases of the Straw Man fallacy, that is, cases in which the arguer mistakenly misrepresents her opponent’s viewpoint, without intending to do so. For example, the desire that my philosophical position is correct may lead me (unintentionally) to focus excessively on the aspects of the opposite account that are seemingly weaker – and, conversely, to neglect aspects that could provide a stronger version of the argument. Hence, I may present an interpretation of my opponent’s viewpoint that exaggerates certain aspects and makes others seem implausible. But this could simply be the result of the unconscious influence that my desire to be right is susceptible to exert on the direction of my attention. After all, as Talisse and Aikin (2006) suggest, the Straw Man fallacy is often the result of a selective interpretation of the opponent’s view, rather than a misrepresentation strictly speaking. On many occasions, they argue, the problem is not that the arguer represents incorrectly the opponent’s argument, but more exactly that she refutes a relatively weak version of that argument.

In this sense, unintentional occurrences of the Straw Man fallacy are somewhat similar to those of the confirmation bias, except that the confirmation bias involves a selective focus on elements that seem to confirm the arguer’s position, whereas the Straw Man fallacy involves a selective focus on elements that seem to falsify the opponent’s position. But the principle at work seems to be the same, namely: the effect of an attentional bias on the arguer’s reasoning. This is due to the fact that the relevance of each element is determined by preconscious processes that can be motivationally biased. As Evans (1995: 147) explains, “people reason with a selective representation of problem information which appears to them to be relevant”.

### 4. Emotional biases and motivated fallacies

---

3 According to Talisse and Aikin (2006), there are in fact two forms of the Straw Man fallacy: the representation form, which corresponds to the standard analysis, and the selection form, which consists in refuting the weakest version of the opponent’s position.

4 To illustrate this aspect, Evans draws a fascinating analogy between the phenomenon of selective relevance in argumentation and the game of chess: “It is well known that experienced chess players look at a position and ‘see’ that a small number of moves (out of many more legal possibilities) need to be considered. Similarly, when analyzing possible replies of the opponent, only a few moves appear relevant” (Evans 1995: 148-149).
The opponents of the motivational account of irrationality often raise the following objection: if emotions are such an important cause of fallacious reasoning, why do they sometimes affect our thinking and other times not? Why don’t people believe systematically whatever they wish to believe? One plausible answer to this question is that “rationality is a matter of degree” (Baron 1988: 36), and that people fall prey or not to irrational thinking depending both on the degree of their emotional attachment to a given standpoint and on the degree of reliability of their methods of reasoning. But there is an additional explanation for the fact that emotions do not always affect significantly the way we reason. As Kunda (1990: 482) points out, this could be due to the fact that “[p]eople do not seem to be at liberty to conclude whatever they want to conclude merely because they want to”. Instead, motivated reasoning often requires a preconscious effort of justification and elaboration without which the desired conclusion would hardly seem plausible to the subject: “There is considerable evidence that people are more likely to arrive at conclusions that they want to arrive at, but their ability to do so is constrained by their ability to construct seemingly reasonable justifications for these conclusions” (Kunda 1990: 480).

This aspect is perhaps most visible in the phenomenon of rationalization, which is characterized precisely by the effort to elaborate reassuring (but false) reasons to justify an attitude or an event. A textbook example of rationalization is the alcoholic who justifies his or her excessive drinking by saying: ‘I am not an alcoholic. I drink because…’ Although the reasons thus provided are never the most likely (the addiction), they often sound reasonable both to the arguer and to her listeners, due to that effort of justification. This is what makes it so difficult to refute the arguments of a person who is rationalizing, particularly if we bear in mind that the mechanism is unconscious and the person is sincerely convinced that those reasons are valid. Furthermore, in many cases the problem isn’t so much that the reasons put forward are false, but more exactly that they are not appropriate to explain the attitude in question. In such cases, the arguer seems to commit the fallacy of misidentifying the cause, either by “falsely identify[ing] X as the cause of Y when on closer inspection a third factor, Z, is the cause of both X and Y”, or by “confus[ing] a cause and effect: identifying X as the cause of Y, when it is actually Y that causes X” (Tindale 2007: 179-180). A typical example of this fallacy is reported by Twerski (1997: 35): Felicia claims that she drinks and uses pills because of the disputes with her husband and because her children treat her with disrespect. Yet, Twerski explains, Felicia is actually confusing the cause and the effect, as it happens so often with addicted persons: “Her husband’s behavior, although unpleasant, is in response to her drinking and pill taking. He cannot communicate meaningfully with her because of her chemical abuse. The children are angry and ashamed that they cannot invite friends over because they fear her unpredictable antics. They have lost respect for her because of her chemical abuse” (Twerski 1997: 35). What is remarkable in this type of case is that the arguer does not fail to take into account the evidence that favors the undesired conclusion; instead, she commits the fallacy of misidentifying the cause and interprets the effects of her drinking as being the reason why she drinks. The undesired conclusion is thus ruled out; at least if we accept the prejudiced premise that the person who drinks for a reason is not an alcoholic.
In other extreme cases of irrationality, such as *self-deception* and *denial*, the anxiety toward the undesired conclusion that \( p \) is so intolerable that the subject fails to infer \( p \) from \( q \), even when \( p \) clearly seems to entail \( q \) (Cohen 1992, Barnes 1997). As Cohen suggests, this is possible because psychological deductions are not as compulsory as logical deductions: “the statement that you believe that \( p \) does not necessarily imply that you believe that \( q \), even where \( q \) is quite a close and well-recognized logical, conceptual, or mathematical consequence of \( p \)” (Cohen 1992: 31). When faced with strong evidence \( q \) in favor of \( p \), the subject’s anxiety that \( p \) seems to inhibit the appropriate inference. The subject acknowledges the evidence, but not its implications. This seems to correspond to a fallacy that some informal logicians call *slothful induction*, that is, “the mistake of underrating the degree of probability with which a conclusion follows from evidence” (Barker 2003: 264). In a sense, slothful induction seems to be the opposite of hasty generalization, given that the later involves ‘jumping to conclusions’ on the basis of insufficient evidence, whereas the former involves failing to draw a conclusion despite the available evidence. For example, in the book *Logical Self-Defense* (1983) Johnson and Blair evoke the case of a woman whose husband worked for an oil company accused of conspiring to fix gasoline prices. In response to the Report of the Bertrand Commission, which made the allegation, the woman wrote:

"Bertrand and the commissioners must be out to lunch. In no possible way could he have one lousy shred of evidence to support their allegations. I can say this because no price fixing occurred, and therefore no evidence for it could exist. My husband has been working for the oil company for 30 years and the company has always been good to him. To say that the industry my husband works for has been ripping off the public for years really irks me (Johnson and Blair 1983: 177)."

As Johnson and Blair suggest, the woman’s argument seems to be biased by her ‘egocentric commitment’: she identifies herself with the oil company’s image, via her identification with her husband, and this attachment causes her to feel extremely anxious about the allegation that the oil companies might have been conspiring to fix gasoline prices: “To say that the industry my husband works for has been ripping off the public for years really irks me”. This anxiety seems to be the motive causes her to defend the oil companies by all means, much to the detriment of the argument’s reasonableness. It is noteworthy that, in doing so, she commits a different fallacy in each statement she advances. First, she commits the fallacy of *slothful induction* by ruling out in advance the very possibility that the evidence might incriminate the companies: “In no possible way could he [Bertrand] have one lousy shred of evidence to support their allegations”. Second, she is clearly *begging the question* when she tries to justify this claim: “I can say this because no price fixing occurred, and therefore no evidence for it could exist”. In fact, if we combine this statement with the previous one, we realize that her reasoning amounts to the following: (1) there is no price fixing because there is no “shred of evidence to support [that] allegation”, and (2) “no evidence for it could exist” because “no price fixing occurred”. And third, she commits the fallacy of *ignoratio elenchi* (or ‘irrelevant conclusion’) when she evokes the fact the “the company has been good to [her husband]”, which has little to with the issue of the fixing of the oil prices.
In the case of wishful thinking, on the other hand, it is not the anxiety that not-\( p \), but rather the desire that \( p \) that seems to affect the subject’s inference. Although the standard description of wishful thinking has the form: ‘I wish that \( p \), therefore \( p \)’, in reality the fallacy is seldom committed in those terms. Few people would directly infer a fact from a desire, claiming for instance ‘I wish to be rich, therefore I am rich’. In most cases, wishful thinking happens without the subject’s awareness of the fact that his belief that \( p \) is motivated by his desire that \( p \). In consequence, the fallacious reasoning at work in wishful thinking is generally more complex. Typically, it takes the form of the argument from consequences: ‘If \( p \) than \( q \). I wish that \( q \). Therefore \( p \)’. But the premise expressing the arguer’s desire often remains implicit, as in the following argument:

Of course I will pass the exam. It’s the only way I can get a job.

The deceiving element in this fallacy is the unjustified assumption that a strong desire to get a job is a sufficient condition to believe that one is capable of doing whatever is necessary to get it. Despite the adage that “If you really want something, you’ll get it”, the fact of the matter seems to be that motivation alone does not guarantee success. In addition, wishful thinking is rooted in a variety of self-serving biases (optimism bias, overconfidence effect, egocentric bias, etc) which tend to reinforce this type of illusion. A striking illustration of such phenomena is the above-average effect (or ‘illusory superiority’), that is, the tendency to overestimate one’s virtues relative to others. According to empirical studies, most people are convinced that they possess better than average competences in many regards. For example, most people consider themselves to be happier, more fair-minded, more skilful behind the wheel and more likely to live past eighty than the average person (Gilovich 1991: 77). Ironically, we are probably even more biased than we think, given that one of the most common illusions is precisely the propensity to believe that we are less biased than the others (Pronin et al. 2004).

5. Conclusion

The purpose of this paper was to examine some of the ways in which goals or motives affect people’s reasoning during a debate. We have seen that motivation affects argumentative reasoning at different levels (selective evidence gathering, selective choice of premises, biased interpretation, and fallacious inference). Furthermore, it was shown that there are strong connections between specific types of motivated biases and specific types of fallacious reasoning, which seems to provide additional support to Kunda’s claim that the ability to arrive at the conclusions we wish to arrive at hinges on the ability to construct seemingly plausible justifications for these conclusions. This sub-intentional effort of justification and argumentation seems to explain (1) the fact that motivated biases give rise to structured (albeit illegitimate) forms of argument, as well as (2) the fact that such fallacious arguments can be so persuasive, both to the arguer (self-deception) and to the audience. But conversely, it would be interesting to investigate in

---

5 Walton (1989: 22-24) distinguishes between two types of ‘argument from consequences’: the argument from positive consequences, where the conclusion is supported by its desirable consequences, and the argument from negative consequences, where the conclusion is supported by its undesirable consequences.
future work what the (il)logical structure of informal fallacies might reveal about the way motivated biases operate.

Although many whistle-blowers alert to the dangers of biased argumentation and ‘manufactured consent’, most analyzes tend to focus on the use and abuse of intentional methods of propaganda. Yet, we have seen that in many cases the arguers are themselves unaware of the fallacies they commit and of the biases that induce them. This means that the sincerity requirement and the care for truth are not sufficient to ensure the rationality of our arguments. Even well-intended arguers might fall prey to motivated illusions and put forward tendentious arguments, given that cognitive biases are unconscious. However, this is not to say that people cannot be held accountable for the irrationality of their arguments: while we are not directly responsible for the fallacies induced by motivated biases, since they are typically unintentional, it seems reasonable to suggest that we are indirectly responsible for creating the conditions necessary to counteract the effect of those biases on our arguments. After all, we have a certain degree of control over the cognitive faculties at work in the process of argumentation. The arguer who is genuinely interested in arriving at an impartial conclusion may, for example, make sure to take into account all the available relevant evidence (when possible). Likewise, one can try to avoid an unintentional misrepresentation of the opponent’s viewpoint (Straw Man) by ‘playing the devil’s advocate’, as Mill (1859: 45) rightly suggests: “If opponents of all-important truths do not exist, it is indispensable to imagine them and supply them with the strongest arguments which the most skilful devil’s advocate can conjure up”. Furthermore, the analytic task of ‘reconstructing’ arguments seems to provide indirectly a reliable method to detect biases and avoid tendentiousness, insofar as it forces the arguer to externalize the argument’s components and to expose its potential weaknesses (Eemeren & Grootendorst 2004). In a broader perspective, it seems paramount to try to develop an ethic of argumentation capable of systematizing these and other strategies of critical thinking in light of a normative account of argumentation.

References


