

University of Windsor

Scholarship at UWindor

OSSA Conference Archive

OSSA 11

May 18th, 9:00 AM - May 21st, 5:00 PM

Don't worry, be gappy! On the unproblematic gappiness of alleged fallacies

Fabio Paglieri

Istituto di Scienze e Tecnologie della Cognizione, CNR Roma

Follow this and additional works at: <https://scholar.uwindsor.ca/ossaarchive>



Part of the [Education Commons](#), [Philosophy Commons](#), and the [Psychology Commons](#)

Paglieri, Fabio, "Don't worry, be gappy! On the unproblematic gappiness of alleged fallacies" (2016). *OSSA Conference Archive*. 119.

<https://scholar.uwindsor.ca/ossaarchive/OSSA11/papersandcommentaries/119>

This Paper is brought to you for free and open access by the Conferences and Conference Proceedings at Scholarship at UWindor. It has been accepted for inclusion in OSSA Conference Archive by an authorized conference organizer of Scholarship at UWindor. For more information, please contact scholarship@uwindsor.ca.

Don't Worry, Be Gappy! On the Unproblematic Gappiness of Alleged Fallacies

FABIO PAGLIERI

Istituto di Scienze e Tecnologie della Cognizione

CNR, Roma, Italy

fabio.paglieri@istc.cnr.it

Abstract. The history of fallacy theory is long, distinguished and, admittedly, checkered. I offer a bird eye view on it, with the aim of contrasting the standard conception of fallacies as attractive and universal errors that are hard to eradicate (section 1) with the contemporary preoccupation with “non-fallacious fallacies”, that is, arguments that fit the bill of one of the traditional fallacies but are actually respectable enough to be used in appropriate contexts (section 2). Godden and Zenker have recently argued that reinterpreting alleged fallacies as non-fallacious arguments requires supplementing the textual material with something else, e.g., probability distributions, pragmatic considerations, dialogical context. Thus fallacies remain *gappy* on all accounts, and this is the hallmark of their failure. However, I argue that such gappiness is typically unproblematic, and thus no more flawed than enthymematic argumentation in general (section 3). This, in turn, calls into question the usefulness of the very notion of fallacy.

Keywords: argumentative theory of reasoning, enthymemes, gappiness, fallacy theory, reasoning errors

1. Introduction: The standard conception of fallacies and its problems

This is how popular wisdom, aka. Wikipedia, defines the notion of fallacy:

A fallacy is the use of *poor, or invalid, reasoning* for the construction of an argument. A fallacious argument may be *deceptive by appearing to be better than it really is*. Some fallacies are committed intentionally to manipulate or persuade by deception, while others are committed unintentionally due to carelessness or ignorance. [my emphasis]¹

Such definition is actually close to the scholarly conception of fallacies, at least as represented by the *Stanford Encyclopedia of Philosophy*:

Two competing conceptions of fallacies are that they are false but popular beliefs and that they are *deceptively bad arguments*. These we may distinguish as the belief and argument conceptions of fallacies. [...] There are yet other conceptions of what fallacies are, but the present inquiry focuses on the argument conception of fallacies. Being able to detect and avoid fallacies has been viewed as a supplement to criteria of good reasoning. The knowledge of them is needed to arm us against *the most enticing missteps we might take with arguments*—so thought not only Aristotle but also the early nineteenth century logicians Richard Whately and John Stuart Mill. [my emphasis] (Hansen, 2015)²

¹ Retrieved September 23 from <https://en.wikipedia.org/wiki/Fallacy>

² Importantly, this entry is authored by Hans V. Hansen, a well-known fallacy scholar.

This view of fallacies as enticing missteps is considered to be the standard conception in the Western tradition (Hitchcock, 2006), and indeed instances of it abound in textbooks and scholarly articles. The following are just two notable examples:

Fallacies are the *attractive nuisances* of argumentation, the *ideal types of improper inference*. They require labels because they are thought to be *common enough* or *important enough* to make the costs of labels worthwhile [my emphasis]. (Scriven, 1987, p. 333)

By definition, a fallacy is a *mistake* in reasoning, a mistake which *occurs with some frequency* in real arguments and which is characteristically *deceptive*. (Govier, 1987, p. 177)

In his recent monograph on fallacies, John Woods (2013) captures the defining features of this view using the acronym EAUI: on the standard conception, fallacies are thought to be *errors* (E) that are *attractive* (A), *universal* (U), and *incurable* (I)—in the sense that people are expected to persevere in fallacious reasoning, even after being shown the error of their ways. After Charles Hamblin’s seminal work on fallacies (1970) revamped the interest in this subject, the EAUI conception have attracted substantial criticism (Massey, 1981; Finocchiaro, 1981; Hintikka, 1987; Hitchcock, 1995; van Eemeren & Grootendorst, 1995; Woods, 2013; Boudry, Paglieri, & Pigliucci, 2015), yet it remains prominent in the literature, as shown by the previous excerpts, and it still inspires the treatment of fallacies in many logic textbooks, where the most common, quick-and-dirty definition of fallacies is “an argument that seems to be valid but it is not”.³

Instead of attempting to cover all the reasons why the EAUI conception is unsatisfactory, here I will only focus on its most glaring difficulties. A *prima face* worry is that this view conflates a normative claim (that fallacies are errors) with various empirical assertions (that they are universal, attractive, and incurable). While this is not automatically a *faux pas*, the fact that such empirical claims appear hard, if not outright impossible to verify should raise a big red flag. How are we supposed to check whether an alleged fallacy is universal, attractive, and incurable? Even if we interpret the universality of these claims as just pertaining a majority of the human population, as charity demands, too many details remain unclear to allow for satisfactory verification. How many exceptions, i.e., people who do not commit a certain error, would disqualify that as a fallacy?⁴ The EAUI conception remains utterly silent on these points. Unfortunately, as Paglieri (2014) argued with respect to Aristotle’s definition of dialectical reasoning, it is never good for a theory to incur a huge empirical debt without having the means to pay it off. This is even truer when defining something which is supposed to be of crucial importance for everyday practice, as it is certainly the case with fallacies.

³ As Maurice Finocchiaro (1981) once pointedly remarked, many textbook cases of fallacies have the habit of existing *only* in textbooks. The most likely reason for this oddity is that when an argument is patently bad (e.g., “If an animal is a gorilla, then it is a primate. X is a primate, therefore X is a gorilla”), people do not find it attractive in the least, hence they do not use it, nor would fall for it if others were to try it.

⁴ Similar worries apply even if one accepts the refined version of universality articulated by Woods (2013, p. 141), according to which fallacies on the standard conception are meant to identify the *typical ways* in which reasoners err, regardless of how frequently they do in fact err. Even *that* claim, however, is far from being proven, as Woods convincingly shows.

Similar considerations recently led Boudry, Paglieri and Pigliucci (2015) to formulate a dilemma in fallacy theory, which they propose to call “the Fallacy Fork”. According to them, if fallacies are construed as demonstrably poor forms of reasoning,⁵ then they have very limited applicability in real life, i.e., few actual instances: thus we get the normative part right (fallacies are indeed error), but we lose completely the empirical features attributed to them by the standard conception—those errors are neither especially attractive nor incorrigible for people, therefore certainly not universal in their occurrence. If, instead, our definitions of fallacies become sophisticated enough to capture real-life complexities, then they can no longer be held up as an effective tool for discriminating good and bad forms of reasoning, since we start talking of fallacious and non-fallacious instances of alleged fallacies—which is exactly what is commonplace in the literature nowadays (for a detailed exposition, see Walton, 1995). In section 2 I will discuss some of these attempts to articulate a theory of “non-fallacious fallacies”, whereas in section 3 I will present evidence supporting the claim that people are not especially prone to obviously mistaken forms of reasoning—at least, not to the extent required by the standard conception of fallacies. For now, let us take stock of the Fallacy Fork’s main implication for the present argument: you can have the E without the AUI or the AUI without the E, but not both at once—hence the standard conception fails to provide a suitable account of fallacies.

Admittedly, the Fallacy Fork is a problem mostly for *structural accounts* of fallacies: that is, theories that describe fallacies as instantiations of flawed reasoning schemes. What makes the Fallacy Fork so pernicious for structural accounts is, quite simply, the abundance of available counterexamples to their proposed definition of a fallacy: arguments that instantiate the allegedly fallacious scheme, yet do not appear flawed, either intuitively or upon reflection. Take the *ad hominem* fallacy as a case in point, i.e., the “error” of accepting or dismissing a claim based on personal features of its source, and consider the following argument: “John’s research is being funded by Marlboro, therefore his conclusions on the lack of negative side-effects of tobacco consumption are unreliable”. The fact that the argument is deductively invalid is as obvious as it is irrelevant: the question is whether it is a bad argument also on weaker, more sensible standards. Here it is obvious that John’s blatant conflict of interest gives excellent reasons to consider his claims as unreliable, insofar as they happen to accord with his personal convenience in the matter. However, the argument fits the structural definition of the *ad hominem* fallacy, even if it is perfectly reasonable to consider it good. Clearly, structural accounts of fallacies strive to remedy this problem by introducing further constraints on their definition of the fallacy in question: so, for instance, Brinton (1995) proposes that *ad hominem* arguments are admissible as long as they are directed at someone’s advocacy of P, while they are fallacious when targeting P itself. Thus our example is admissible because the conclusion refers to John’s advocacy (his claims are described as unreliable), whereas the argument would be fallacious if the conclusion pertained the facts of the matter (for instance, his claims were considered as false). This makes perfect sense in principle, but now the account is exposed to the second horn of the Fallacy Fork: many arguments have a partly testimonial character, and stand or fall with the credentials of the person advocating them. In our example, in practice considering John’s claims to be unreliable is

⁵ Crucially, “demonstrably poor” is much broader than “deductively invalid”, and deliberately so. The fact that many perfectly sensible arguments are not deductively valid has now been remarked endlessly in the literature, so we should stop flogging the deductive horse—it is long dead. The problem of fallacy theory is a different one: namely, the fact that, when an argument is bad even according to non-deductive standards, laypeople are still smart enough to avoid buying it, thus making the alleged deceptiveness of fallacies something of a myth.

tantamount to consider them as probably false, since they end up being dismissed either way. Thus the notion of fallacy, once properly refined to avoid mislabeling good reasoning, ends up losing much of its practical significance.

However, fallacies have also been defined in non-structural terms, but rather *dialectically*, as infractions of some norms of proper dialectical engagement (Hamblin, 1970; van Eemeren & Grootendorst, 1995; Walton & Krabbe, 1995).⁶ These approaches manage to provide a solid theoretical grasp on fallacies without incurring in the Fallacy Fork, yet they suffer instead from a labelling problem. In fact, even if the same names are used, it is hard to see how the original fallacy may constitute an instantiation of its dialectical definition: as Woods (2013) noted with respect to pragma-dialectics (PD),

the striking thing of the PD approach to fallacies is that whereas it show no reluctance to forgo the traditional concept of fallacy, it hangs on tightly to the traditional *list*. [...] This is puzzling, [because] with the possible exception of begging the question [...], which PD theorists may see as a violation of the unexpressed premise rule, none of the [traditional fallacies] is in any obvious way the violation of a PD rule. What PD rule does many questions violate? What does hasty generalization violate? (pp. 507-508)

Thus, while in general I find dialectical theories of fallacies more viable than structural ones, I doubt these two approaches really focus on the same phenomena, and in what follows I will limit my analysis to a structural understanding of fallacies—which of course can be informed by dialectical considerations, yet sees fallacies as reasoning errors, rather than missteps in communication.

2. The quest for non-fallacious fallacies

Taken literally, the expression “non-fallacious fallacies” is clearly oxymoronic, in that it refers to an error that is not an error. This is not the sense in which I will use it here, nor in which it is used elsewhere: instead, non-fallacious fallacies stands as shorthand for “non-fallacious arguments that structurally match the definition of a certain fallacy”. The adoption of this weird label is intended to emphasize the problems inherent in clinging to the term “fallacy”, while dealing with patterns of reasoning. Even if argumentation scholars nowadays are fully aware that virtually every type of informal fallacy has its legitimate counterpart, the label “fallacy” is still very much in use, as discussed. While this habit is not logically contradictory, it is indeed conceptually confusing and rhetorically pointless: so, by the end of this chapter, I hope I will have started giving reasons to discontinue the practice, and send the term “fallacy” into retirement.

In the meantime, let us review some other examples of non-fallacious fallacies, the weird beast that has preoccupied fallacy theorists over the last 40 years or so. My first specimen is based on the *ad ignorantiam* fallacy, that is, drawing a conclusion based on absence of evidence

⁶ This distinction between structural and dialectical accounts of fallacies is reminiscent of Walton’s (1995) articulation of fallacies in two broad categories: *paralogisms* (fallacious premise-inference structures that fail to meet some necessary requirement of an argumentation scheme) and *sophisms* (fallacious tactics used to unfairly try to get the best of a speech partner in the course of a dialogical exchange).

to the contrary. As the old epistemological maxim goes, “absence of evidence is not evidence of absence”. But now consider the two following arguments:

- My wife will not give me a present for my birthday, since I have received no indication to the contrary from her.
- My flight tomorrow from Schiphol will not be delayed, since I have received no indication to the contrary from the airline.

Although structurally both arguments instantiate the *ad ignorantiam*, they elicit very different intuitive appraisals: the first one is immediately perceived as extremely weak, if not outright erroneous, whereas the second will be good enough for anyone except the most skeptical person—in fact, we routinely reason that way the day before taking a trip. Nor is the difference particularly hard to account for: the presumptive weight of the absence of evidence for X depends on whether or not there is a justified expectation to obtain that evidence, whenever X is indeed the case. If that expectation is low or non-existent, then absence of evidence carries very little presumptive weight; but if the expectation is robust enough, then we are presumptively justified in concluding that X is unlikely, given that lack of evidence. This is precisely what supports our intuitions regarding these two arguments: whereas I have no reason to expect my wife to broadcast her plans for my birthday present (indeed, I have reasons to assume the opposite), I know as a fact that most airlines will make every possible effort to keep their customers well informed of any schedule change. This is why absence of evidence does suggest evidence of absence in the second case, but not in the first. Crucially, however, you do not need to appreciate the reason behind the difference to be able to detect it: laypeople would be perfectly capable of correctly rating the strength of the two arguments, even if they were unable to articulate a justification for it (empirical evidence to that effect will be reviewed in section 3, in case someone doubts it).

Nor is this phenomenon specific of the *ad ignorantiam*. If we now move to consider the *ad verecundiam*, i.e., accepting a certain conclusion on the say-so of some source, the same pattern arises. Consider the following two examples:

- Hollywood celebrities claim that homeopathy is effective, so it is likely to be effective.
- Medical practitioners claim that homeopathy is effective, so it is likely to be effective.

Again, the structure of both arguments is the same and fits the description of the *ad verecundiam* fallacy: however, the latter is patently stronger than the former, as any competent speaker immediately intuits. There is of course no big mystery behind such an intuition: here the different appraisal of the arguments mostly depends on considerations of expertise. While there is absolutely no reason to defer to Hollywood celebrities as experts on homeopathy, the competence of medical practitioners on this matter is manifest. Moreover, medical practitioners are typically assumed to be skeptical of homeopathic remedies, thereby making their endorsement of such remedies all the more convincing. Conversely, Hollywood celebrities are often thought as having a penchant for outlandish medical treatments, thus making their claims

on homeopathy less likely to carry presumptive weights to people that are not already persuaded of its merits.

At this point, a defender of fallacy theory may consider retreating to a more entrenched position: that is, s/he may be willing to concede that non-fallacious fallacies abounds when it comes to informal fallacies, yet insists that no such thing exists, when it comes to formal fallacies, the real “bad boys” of logical textbooks. The prime examples of formal fallacies are, of course, *denying the antecedent* (DA) and *affirming the consequent* (AC). Given a conditional statement such as “If a bird is a flamingo, then it is pink”, there are two correct reasoning patterns one can apply to it: *modus ponens* (from the conditional and the fact that “That bird is a flamingo”, it is derived that “That bird is pink”) and *modus tollens* (from the conditional and the fact that “That bird is not pink”, it is derived that “That bird is not a flamingo”). DA and AC are their fallacious cousins: DA uses the negation of the antecedent to erroneously infer the consequent (from the conditional and the fact that “That bird is not a flamingo”, it is derived that “That bird is not pink”), whereas AC mistakes the truth of the consequent as evidence for the truth of the antecedent (from the conditional and the fact that “That bird is pink”, it is derived that “That bird is a flamingo”).

The fact that DA and AC are deductively invalid is as obvious as it is uninformative. The real question, instead, is whether these reasoning patterns are unequivocal mistakes under *any* legitimate standard of inference, besides deductive validity. The short, non-technical answer to that question is that it depends on how many non-flamingo pink birds are out there—more generally, it depends on the likelihood of alternative reasons for the consequent, other than the antecedent. As Luciano Floridi (2009) pointed out, DA and AC can be coherently interpreted as “Bayesian ‘quick and dirty’ informational shortcuts [which] assume [...] that there are no false positives (double implication), or that, if there are, they are so improbable as to be disregardable (degraded Bayes’ theorem)” (p. 322; see also Stone, 2012). Take the following reasoning pattern: “When I have the flu, I have a fever. Today I have a fever, therefore I have the flu”. This is certainly not deductively valid, but is it also blatantly mistaken on any other legitimate standard of inference? Ultimately, whether the argument is a fallacious instance of AC or an acceptable inference to the best explanation depends on the likelihood of alternative reasons for the observed symptoms, i.e., my fever: if this likelihood is low enough, in relation to the conditional probability of having a fever given the flu, the inference goes through without particular problems. Granted, it is still defeasible, and necessarily so, since no amount of probabilistic considerations can produce deductive validity out of thin air. But defeasibility, in and by itself, is not enough to brand it as rubbish on every normative standards. The same applies to our flamingo example: in a world where non-flamingo pink birds are rare enough, then for a bird being pink is a reliable (albeit defeasible) indicator of it being a flamingo (AC), whereas not being a flamingo drastically reduces the likelihood for that bird to be pink (DA).

More to the point, not only non-fallacious formal fallacies are possible, they might even be as widespread as non-fallacious informal fallacies. Compare the following:

- If an animal is a chimpanzee, then it is a primate. King Kong is a primate, therefore King Kong is a chimpanzee.
- When I have the flu, I have a fever. Today I have a fever, therefore I have the flu.

As usual, both arguments match the structural blueprint of a fallacy (AC, in this case): yet the first one is patent nonsense, whereas the second one sounds good enough under the appropriate circumstances, e.g., for a first, rough-and-ready diagnose of one's own medical condition. What matters, however, is the fact that their difference in strength is once again intuitively accessible to any competent speaker. Who in his/her right mind would fall for (or even use in the first place) the King Kong argument? In contrast, making an educated guess on a certain illness based on one of the symptoms that illness causes is a perfectly legitimate, albeit fallible, epistemic practice, assuming the probabilities provide enough support to justify such inference.

Now, of course the definition of the various fallacies, formal and informal, can be made precise enough to exclude non-fallacious instances: indeed, much of the recent efforts in fallacy theory have been directed to detail the conditions under which a certain argumentative pattern, previously labeled as fallacious in general, can be used legitimately, and when instead the accusation of fallaciousness still stands.⁷ Yet comparing generic vs. specific definitions of fallacy brings us again into the grip of the Fallacy Fork. If fallacies are defined precisely, then they instantiate reasoning mistakes that are, however, not particularly common in everyday practice, nor especially attractive. If, on the other hand, fallacies are defined generically, then most of their instances turn out to be perfectly legitimate. Since the intuitive notion of a fallacy is meant to capture a mistake that occurs with alarming frequency, either way the expression fails to deliver on that intuition: the precise approach captures mistakes that are relatively infrequent, whereas the generic view encompasses argumentative patterns that are widespread but typically fine. Thus on both accounts the term fallacy is being misused, with no good reason to do so, other than a nod to historical tradition. Precise “fallacy” theories deal with real reasoning mistakes that are neither particularly frequent nor very appealing to arguers, hence should not be named “fallacies”. And generic “fallacy” theory does not even identify mistakes, so the term “fallacies” is out of place there as well.

3. The gappiness of fallacies

There is, however, a way of trying to defend the legitimacy of the notion of fallacy in something akin to its standard conception. A brilliant effort in that direction has recently been made by David Godden and Frank Zenker (2015): these authors first provide a comprehensive and well researched review of recent attempts to show that formal fallacies, i.e., DA and AC, are ordinarily cogent—that is, “well-reasoned [on] a generic, theoretically-neutral, objective, normative standard of argumentative or inferential goodness” (p. 89). Then they argue *against* this position, by showing how, on all these attempts, “the acceptability of the conclusion of DA and AC arguments depends on factors not asserted by the stated conditional” (Godden & Zenker, 2015, p. 89). Godden and Zenker (2015) think that such gappiness is what makes AC and DA problematic, since on their view AC and DA “fail to be cogent whenever they conspicuously fail to cite as reasons the conditions on which the acceptability of their conclusions properly depends” (p. 89).

⁷ Douglas Walton, for instance, authored a series of monographs on how to distinguish truly fallacious from spuriously fallacious uses of argument patterns traditionally considered to be fallacies in general: among others, *ad verecundiam* (1997), *ad hominem* (1998), *ad ignorantiam* (1999), *ad baculum* (2000). For similar efforts by other scholars, see Mackenzie (1980), Massey (1981), Brinton (1995), Goodwin (1998), Woods (1998), Levi (1999), Hitchcock (2007), and Mizrahi (2010), among many others.

Before criticizing their line of argument, I want to stress something on which we are in full agreement: all recent attempts to show AC and DA to be cogent (non-fallacious, in the sense I have been using it here) hinge on supplementing the interpretation of the stated conditional with something not included in its explicit formulation. In other words, Godden and Zenker (2015) are correct in emphasizing the gappy nature of formal fallacies—indeed, I think such gappiness is characteristic of *any* so called fallacy, formal or informal, even though their paper only focuses on AC and DA. What exactly is supposed to fill the gap depends on the theoretical approach being considered: on a Bayesian reading of formal fallacies (Korb, 2004; Hahn & Oaksford, 2007; Floridi, 2009; Stone, 2012; Hahn, Oaksford, & Harris, 2013; Zenker, 2013), the gap would be filled by probabilistic considerations; on a reconstruction of the inference in terms of Stalnaker’s (1968) conditionals, cogency would depend on similarity relations among possible worlds; on interpretations that invoke conditional perfection⁸ as a way of justifying the rationality of AC and DA (Geis, & Zwicky, 1971; Moldovan, 2009), what matters would be pragmatic indicators; and on dialectical theories of fallacies (Hamblin, 1970; van Eemeren & Grootendorst, 1995; Walton & Krabbe, 1995; Godden & Walton, 2004), filling the gap would require considering the contextual features of the dialogue. In spite of the key differences among these alternative approaches, it is always the case that the cogency of the alleged fallacy depends on something that is not explicitly mentioned in its utterance.

The crux of the matter, however, is whether or not this gappiness constitutes a significant flaw. Godden and Zenker (2015) think so, whereas I beg to differ. My argument against their conclusion can be summarized as follows:

1. Gappiness per se cannot be considered a capital sin, lest we throw away the baby of enthymematic reasoning with the bathwater of alleged fallacies.
2. What would make the gappiness of fallacies “bad” are the same criteria that would disqualify as illegitimate any other enthymeme: namely, that what is left implicit is controversial or unclear.
3. Thus whether the gappiness of fallacies is problematic or not ultimately rests on an empirical question, to wit, whether the “missing ingredient” in non-fallacious fallacies is typically controversial or unclear.
4. Extant evidence supports the claim that most alleged fallacies are gappy in unproblematic ways, vindicating the intuition that there is nothing wrong with non-fallacious fallacies, and thus making the continued use of the label “fallacy” problematic, for the reasons exposed above.

Point 1 should be rather self-evident: gappiness is a property of another well-known family of arguments, i.e., enthymemes, and there is an abundance of different theories on how their missing premise is to be reconstructed (Ennis, 1982; van Eemeren & Grootendorst, 1982, 1983; Gerritsen, 2001; Walton, 2001, 2008; Paglieri & Woods, 2011); even approaches that deny the standard characterization of enthymemes as argument with a missing premise (Hitchcock,

⁸ Conditional perfection refers to the interpretative tendency to “perfect” a stated conditional to an unstated biconditional, due to pragmatic considerations. The standard examples involve promises or threats, e.g., a statement like “If you mow the lawn, I’ll give you five dollars” is typically (and correctly) interpreted as shorthand for “If *and only if* you mow the lawn, I’ll give you five dollars”. This interpretation, clearly, makes alleged instances of AC and DA utterly unproblematic, since they are just applications of MP and MT to the other direction of the biconditional.

1985; 1998) make their validity, or lack thereof, dependent on something which is not explicitly stated in the argument itself, e.g., a Toulminian warrant (Toulmin, 1958), which would still count as a form of gappiness in the sense implied by Godden and Zenker (2015). However, nobody thinks that all enthymemes are automatically flawed due to their gappy nature: indeed, the whole point of their analysis is to find ways of (a) interpreting their structure to (b) assess how good or bad they are *qua* arguments. If gappiness was enough of a fault to consider all gappy arguments fallacious, then all enthymemes would have to be considered fallacious; inasmuch as we deny the latter, we are also denying that gappiness per se is problematic. In fact, *prima facie* the gappiness of fallacies seems analogous to the gappiness of enthymemes, thus whether it is problematic or not will depend on the same features that matter for enthymemes: namely, whether what is left unstated is controversial or unclear (point 2). This is of course an empirical question, so I take point 3 to be also self-evident. As for how to respond to this empirical question, the experimental evidence I will shortly review suggests that the gappiness of fallacies is typically unproblematic (point 4).

Before discussing the experimental data, however, it is worth looking at how Godden and Zenker (2015) reply to this line of criticism in their paper (pp. 121-126).⁹ Basically, they concede points 1 and 3, while disputing 2 and 4. As I see it, their defense is two-pronged (the labels are mine):

- *The what argument* (contra 2): the flaw of fallacies is not gappiness per se, but rather *what* they are gappy about—in other words, their gappiness is different, and *worse*, than the gappiness of ordinary enthymemes.
- *The bias argument* (contra 4): evidence on how good people are at discriminating good vs. bad “fallacies” is undermined by results on *other reasoning biases* that plague our cognitive processes.

Let us first consider the what argument. This is how Godden and Zenker (2015) put it:

We do not claim that the incogency of DA and AC arguments results from *that* something has been left unstated; rather the problem we identify concerns *what* has been left unstated. [...] DA and AC arguments fail to assert the conditions on which the truth of their conclusions, and indeed the positive relevance of their stated premises, depend. These are properly construed as reasons, not background assumptions. Premises that are unproblematically supplemented to putatively enthymematic arguments are, minimally, ones that are *reasonably acceptable* to both arguer and audience, and also *accepted by* both arguer and audience. Only then can one take it for granted that these premises are not at issue—that they *go without saying*. Yet, with DA and AC arguments this does not seem to hold. The unstated conditions [...], on which the cogency of DA and AC arguments depend, do not ordinarily go without saying. [...] Indeed, because they give the very

⁹ Full disclosure: I was one of the anonymous reviewers of their paper, whose comments prompted Godden and Zenker (2015) to include a rejoinder in the published article. Thus the present chapter should be seen as the public continuation of a debate started in private. As already mentioned, the fact that I still disagree with their position on DA and AC arguments does not subtract in any way from the value of their paper, especially as an excellent overview of various attempts to articulate the cogency of these alleged fallacies.

conditions on which the cogency of the given argument depends, they are precisely the kinds of claims that are, or should be, at issue, and should therefore be expected to be found among the stated premises of the argument [emphasis in original]. (pp. 122-123)

The claim here is that what is missing in DA and AC arguments is qualitatively different from what is missing in ordinary enthymemes, and the crux of the difference is that the former fail to assert the conditions on which the truth of the conclusion and the relevance of the stated premises depend, whereas the latter are not supposed to do the same. Additionally, Godden and Zenker also claim that what is missing in DA and AC arguments does not typically go without saying, because it is either not reasonably acceptable by all parties, not actually accepted by them, or both. Unfortunately, I fail to see how textual evidence would support either of these claims. Compare the two following arguments:

- When I have the flu, I have a fever. Today I have a fever, therefore I have the flu.
- Socrates is a man, therefore he is mortal.

The former is a potentially acceptable instance of an AC argument, depending on the usual relevant assumptions on probabilities, whereas the latter is “the father of all enthymemes”, i.e., the most frequently cited example of a valid enthymematic argument. What is missing in the first case is, roughly, (A) “The likelihood of alternative causes to my fever is low enough”, whereas what is missing in the second case is something like (B), “All men are mortals”. Now, as anyone can easily verify, *both* (A) and (B) support their respective argument in that they “assert the conditions on which the truth of their conclusions, and indeed the positive relevance of their stated premises, depend”, as Godden and Zenker (2015) put it. In particular, if (B) was not the case, i.e. some men happened to be immortal, then the truth of the conclusion “Socrates is mortal” would no longer be established, and the relevance of the stated premise would also be in question—a lot or a little, depending on the percentage of immortal men over the whole population. On the face of it, the truth of the conclusion and the relevance of the stated premise depend on the missing element as much in enthymemes as they do in fallacies, contra Godden and Zenker. Besides, filling the gap appears to be equally unproblematic in both cases (against, contra Godden and Zenker): the fact that someone claiming to have an illness based on one of its symptoms is assuming other explanations to be unlikely, or at least significantly less likely than the proposed diagnosis, “goes without saying” as much as the reconstruction of the major premise in the Socrates enthymeme. Of course, the *truth* of the implicit assumption is much more questionable in the first case (i.e., alternative causes for a fever are unlikely) than in the second one (i.e., all men are mortal), based on background knowledge. But this has nothing to do with correctly *interpreting* the arguments, which is all that matters for the present discussion: given the enthymeme “Dumbo is an elephant, therefore he flies”, we all interpret it (correctly) as relying on something like “All elephants fly”, even if we know this assumption to be false. But as far as correct interpretation is concerned, what is missing is equally important, and equally unproblematic for arguers to individuate, for both enthymemes and alleged fallacies.

Now it is time to turn to the experimental record. I will first summarize some key findings on fallacy interpretation, and then discuss the results that, according to the bias

argument proposed by Godden and Zenker (2015), would call those findings into question. Ulrike Hahn and Mike Oaksford (2006; 2007), among others, have demonstrated in recent years that people are intuitively good at discriminating between strong and weak instances of the same, potentially fallacious argument schemes, and they do so in ways consistent with a bona fide inference standard (Bayesian update) and showing sensitivity to rationally relevant factors, such as prior belief, argument strength, and nature of the evidence (see also Korb, 2004; Oaksford & Hahn, 2004; Hahn, Harris, & Corner, 2009; Corner & Hahn, 2012; Harris, Hsu, & Madsen, 2012; Hahn, Oaksford, & Harris, 2012, 2013; Collins, Hahn, von Gerber, & Olsson, 2015). Since reviewing this still growing body of literature is not my aim here, I will just outline the gist of their methodology and findings with respect to a specific fallacy, i.e., *ad ignorantiam*, for the sake of illustration. Nonetheless, it is worth emphasizing that the same approach was applied with similar results to a garden variety of other alleged fallacies (e.g., *ad hominem*, *ad populum*, *petitio principii*, and *slippery slope*) and that preliminary theoretical considerations suggest that this kind of Bayesian explanation may be extended to virtually all the fallacies in the classic catalogue (Hahn & Oaksford, 2006).

Let us focus now on the *ad ignorantiam*. To study people's intuitive appraisal of this type of argument, participants are first presented with some ground conditional that they are asked to accept as valid, e.g., "If Drug A produces toxic effects in legitimate tests, then Drug A is toxic". Then the first manipulation is introduced, in terms of *argument type* (positive vs. negative), as follows:

- Drug A is toxic because a toxic effect was observed (positive argument).
- Drug A is not toxic because no toxic effects were observed (negative argument).

On a deductivist reading, the first argument is a valid instance of *modus ponens*, whereas the second argument is a blatant case of DA. Even from a non-deductivist perspective, the latter could still be regarded as instantiating a fallacious *ad ignorantiam*. Ordinary speakers, however, have different intuitions: whereas they typically consider the first argument to be the stronger one, they do not immediately regard the second one as flawed, but rather as a weaker yet potentially valuable inference, depending on other factors. One of these factors is the *amount of evidence* being considered, which can be easily manipulated as follows:

- Drug A is not toxic because no toxic effects were observed in one test (weak evidence).
- Drug A is not toxic because no toxic effects were observed in fifty tests (strong evidence).

The latter argument is clearly stronger than the former, and participants have no difficulty in intuiting the difference. The same is true for another factor that affects the credibility of the conclusion, namely, *prior belief* in it. This is manipulated by using short vignettes, like the following:

Barbara: Are you taking digesterole for it?
Adam: Yes, why?
Barbara: Well, because I *strongly* believe that it does have side effects.
Adam: It does have side effects.
Barbara: How do you know?
Adam: Because I know of one experiment in which they found side effects.

Now these vignettes can be used to manipulate all three factors at once (argument type, amount of evidence, and prior belief) and see how they affect people's intuitions. The previous vignette represents the combination of a positive argument with weak evidence and strong prior belief. The following instead conveys the opposite combination, i.e., a negative argument with strong evidence and weak prior belief (all other combinations are easy to obtain, of course):

Barbara: Are you taking digesterole for it?
Adam: Yes, why?
Barbara: Well, because I *weakly* believe that it does not have side effects.
Adam: It does not have side effects.
Barbara: How do you know?
Adam: Because I know of fifty experiment in which no side effects were found.

When participants are presented with similar vignettes and asked to what extent Barbara will accept Adam's argument, their responses (i) show sensibility to all three factors (ii) in ways that match almost perfectly the Bayesian prediction (Oaksford & Hahn, 2004; Hahn & Oaksford, 2007). Thus the relevant finding is not just that people have consistent intuitions on how to discriminate between actual fallacies and their non-fallacious counterparts, but also that such intuitions are grounded on a legitimate normative standard, namely, Bayesian update. Similar results, which have been established for a variety of alleged fallacies, lead me to claim that "filling the gap" in these instances is no more problematic than in any other case of enthymematic reasoning, contra Godden and Zenker's (2015) what argument.

However, this is where their bias argument is supposed to kick in, casting doubt on the relevance of the empirical findings I just briefly reviewed. This is how Godden and Zenker (2015) present their case:

One cannot simply assume by fiat that arguers sensitively track the unstated cogency conditions on which DA and AC inferences depend. The empirical evidence purporting to show that arguers do so is, in our view, equivocal and inconclusive. Moreover, other empirical evidence suggests that arguers instead (or perhaps also) attend to various irrelevant aspects of reasoning problems, while neglecting relevant aspects. For example, studies of two-premise conditional reasoning show that reasoners respond to logically irrelevant aspects such as premise order (Giroto, Mazzocco and Tasso, 1997) and to whether antecedent or consequent conditions are negated (Evans and Lynch, 1973), resulting in the matching bias (Evans, 1998). Yet other studies show that reasoners fail to attend to logically relevant information. For example, one of the first cognitive biases to be named, the confirmation bias (Wason, 1966), regularly registers as being alive and well (Nickerson, 1998; Mendela et al., 2011). Similarly, base-rate neglect

(Eddy, 1982) (which can be mitigated, though not eliminated, by presenting relevant information in a frequentist format (Gigerenzer and Hoffrage, 1995)) remains prevalent and recalcitrant (Barbey and Sloman, 2007). (pp. 124-125)¹⁰

It is unfortunate that Godden and Zenker do not elaborate on their claim that the evidence in question is “equivocal and inconclusive”: without some indications on why they consider this to be the case, it is hard to rebut the accusation in a meaningful way—especially since, as far as I can see, the evidence is not equivocal at all, but rather quite convergent on the point in question. As for it being inconclusive, well, it is certainly open to falsification, as any form of (good) empirical evidence is; yet it is not *especially* inconclusive, so I fail to see the relevance of this remark. Perhaps Godden and Zenker want to suggest that what makes this evidence problematic is the existence of those other reasoning biases they mention in this passage, i.e., matching bias, confirmation bias, and base rate neglect. If that is the case, then the point of course is whether or not such biases in fact undermine the quality of our folk intuitions on fallacies. Again, no argument is provided to support such conclusion, so perhaps Godden and Zenker think the facts speak for themselves. But when I started browsing some of the papers they mention, I found nothing speaking against the validity and robustness of the findings gathered by Hahn, Oaksford, and others.

By way of example, consider the paper on premise order by Girotto, Mazzocco and Tasso (1997): this presents a series of studies aimed at corroborating Johnson Laird’s mental model theory of conditionals, and 6 out of the 8 studies deal only with valid forms of inference (MP and MT), whereas the last two studies conclude that “conditional fallacies [i.e., AC and DA] are *not* significantly affected by the premise order” [my emphasis] (Girotto, Mazzocco, & Tasso, 1997, p. 1). Far from proving that ordering effects undermine our ability to evaluate potentially fallacious reasoning, this study shows that people are relatively immune from such effects when it comes to deal with possible fallacies—thus actually scoring a point *against* Godden and Zenker’s claim. Or consider the matching bias, i.e., “a tendency to see cases as relevant in logical reasoning tasks when the lexical content of a case matches that of a propositional rule, normally a conditional, which applies to that case” (Evans, 1998, p. 45). Godden and Zenker (2015) cite a paper by Evans and Lynch (1973) as presenting evidence that would lead us to doubt people’s folk intuitions on fallacy discrimination. However, the paper in question deals with a rather different topic: Evans and Lynch introduce a modification of the classic Wason’s selection task, thus showing that the typical mistakes observed in this task do not depend on a tendency to look for confirmation in evaluating conditional statements (Wason’s original hypothesis), but rather on the matching bias itself. While this is certainly interesting, it also suggests that these mistakes are superficial rather than logical: as Evans and Lynch (1973) note, “the overall pattern, with matching bias cancelled out, gave no evidence for a verification bias, indicating instead that the logically correct values were most frequently chosen” (p. 391). How is this relevant for establishing the value of our folk intuitions on fallacy discrimination? No obvious answer is forthcoming, since some alleged fallacy could be explained by invoking a matching bias (e.g., AC), whereas others could not (e.g., DA). More to the point, however, is the fact that this bias does not seem to have any import for our ability to discriminate between, say, a truly fallacious instance of AC and a *prima facie* reasonable inference to the best explanation, since in both cases

¹⁰ To avoid clogging the bibliography of this paper with unnecessary references, I have included in it only the articles cited by Godden and Zenker that are subsequently discussed in the present article. For all other references mentioned in this passage, I refer interested readers to the original essay by Godden and Zenker (2015).

the minor premise (the affirmation of the consequent) would “match”, in Evans’ sense, the original conditional (major premise).

In light of these considerations, I regard the bias argument as being both underdeveloped, thus hard to assess properly and reply to adequately, and based on empirical evidence that has no clear bearing on the matter under discussion, that is, whether or not our folk intuitions on fallacy discrimination are reliable enough. As such, it does not suffice to persuade me that the gappiness of fallacies is especially problematic, over and above the gappiness of enthymemes in general. Moreover, the claim that well-known biases of individual reasoning are worrisome in the dialogic context where alleged fallacies typically occur and are evaluated is at odds with the *argumentative theory of reasoning* recently proposed by Hugo Mercier and Dan Sperber (2011), which suggests the opposite—biases are expected to be particularly widespread in solo reasoning, whereas they largely disappear when the appropriate social context is present. According to Mercier and Sperber (2011), reasoning did not evolve as a correction mechanism for mistaken intuitions or as a support system for individual decisions, contrary to the view championed by Kahneman and others (e.g., Kahneman, 2011), but rather as a tool to effectively engage in a special type of social activity—namely, argumentation. On this view, the function of reasoning is to find and evaluate reasons in dialogic contexts, not through solitary introspection. The evolutionary rationale of the theory has been extensively discussed elsewhere (Sperber et al., 2010; Mercier & Sperber, 2009, 2011; Mercier, 2013) and it is of secondary importance here. What matters is that the argumentative theory of reasoning offers a strikingly different interpretation of the findings mentioned by Godden and Zenker (2015). Mercier and collaborators argue that, if the function of reasoning was indeed to correct mistaken intuitions, then the high incidence of errors observed in laboratory tasks would be hard to account for, especially considering tasks in which reaching the correct solution only requires the application of basic knowledge in logic (Wason, 1966), statistics (Tversky & Kahneman, 1982), probability theory (Tversky & Kahneman, 1983), or mathematics (Frederick, 2005). Moreover, according to Mercier and colleagues the reason why people on their own fail to achieve better performance in such tasks is not because they try to reason correctly and fail, but rather because reasoning itself leads them further into error, by strengthening, rather than questioning, their initial, mistaken intuitions, in line with the literature on motivated reasoning (Kunda, 1990).

In contrast, the very same features that make solitary reasoning so desperately flawed, e.g., myside bias (Nickerson, 1998; Stanovich & West, 2007; Mercier, 2010) and reasoning laziness (Kuhn, 1991; Perkins, Farady, & Bushey, 1991; Kahneman, 2011), make perfect sense in a social context, where unilaterally making your case as strong as possible is (i) personally advantageous and (ii) collectively counterbalanced by the critical scrutiny others will direct against your claims, so that (iii) directing your cognitive efforts towards the evaluation of the arguments of others, rather than your own, is just a sensible allocation of your limited resources, provided this laziness is selective, i.e., it affects self-evaluations and not the cross-examination of the arguments made by others (as shown in Trouche, Johansson, Hall, & Mercier, forthcoming). In support of this interpretation, Mercier and collaborators cite the significant benefits of group discussion on performance in a variety of tasks that individuals fail to master alone, such as the Wason selection task (Moshman & Geil, 1998), in which group performance was on average four times higher than individual results (data analyzed in Mercier, Trouche, Yama, Heintz, & Giroto, 2015). Recently, similar results were reported also with respect to the Cognitive

Reflection Task,¹¹ where simply exposing participants to each other solutions, with no opportunity for actual dialogue among them, was sufficient to rapidly converge on the correct answer (Rahwan, Krasnoshtan, Shariff, & Bonnefon, 2014). More generally, the benefits of group discussion have been observed for “problems or decisions for which there exists a demonstrably correct answer within a verbal or mathematical conceptual system” (Laughlin & Ellis, 1986, p. 177). The relevant fact seems to be the ability of the correct answer to assert itself: as long as one participant gets it right, other members converge on that solution, even if it is originally held only by a minority, or even by a single individual, and independently from how confident the original “truth-bearer” is (Trouche, Sander, & Mercier, 2014). The fact that people, including experts, systematically underestimate the benefits of group discussion on reasoning performance (Mercier et al., 2015) is unfortunate and contributes to explain the widespread picture of laypeople as “bad arguers”, but it does not alter what the evidence suggests: provided with the appropriate social context in which argumentation is meant to take place, we do not argue (nor reason) nearly as poorly as the standard interpretation of so called “fallacies” would lead us to believe.

Finally, it is worth noting that argumentative fallacies and psychological biases are different classes of phenomena: while it may be productive to study their interplay (and I believe it is), the two labels should not be used as synonymous.¹² To be fair, I am quite sure Godden and Zenker (2015) would wholeheartedly subscribe to this sentiment, and I do not intend to accuse them of confusing these two concepts in their paper. Yet the fact that data on a garden variety of psychological biases are mentioned as being automatically relevant for the ongoing debate on argumentative fallacies, with no further reason offered, does raise a red flag, so I think a minor cautionary note is in order. To put it simply for the sake of brevity, a bias in psychology is defined as a cognitive heuristic, typically automatic and inflexible, that may or may not lead to error. In fact, the most heated debate on this subject hinges on whether biases should be regarded mostly as carriers of irrationality (in a nutshell, the position held by Kahneman and Tversky; e.g., Tversky & Kahneman, 1982, 1983; Kahneman, 2011), or rather as “simple heuristics that make us smart”, i.e., useful strategies in our adaptive toolbox (the so called ecological rationality approach, championed by Gigerenzer and others (e.g., Gigerenzer & Selten, 2001; Gigerenzer, Hertwig, & Pachur, 2011)). In contrast, a fallacy in argumentation theory is a mistake by definition, except when the term is used as shorthand for “something that looks like a fallacy but may be not”, as in “non-fallacious fallacies”—for my reservations on this strange usage, see the end of section 2. At the same time, fallacies are not necessarily considered heuristic, and

¹¹ In the Cognitive Reflection Task (CRT from now on; see Frederick, 2005), participants are presented with three simple problems that elicit a strong, intuitive, yet mistaken response, which prevents most people from reasoning out the correct solution—even if such solution is well within their cognitive powers. Performance on the CRT is notoriously low even in highly educated and intellectually brilliant samples (see data in Frederick, 2005), and it is taken to measure the ability to inhibit an intuitive mistake, rather than general intelligence. By way of example, one of the three problems in the CRT is the following: “A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?” Here the intuitive mistake is 10 c, whereas the correct response is 5 c.

¹² Matters are complicated by the fact that psychologists have appropriated the concept of fallacy and used it in their own work as, indeed, roughly equivalent to bias (e.g., Pohl, 2004). As a result, some well-known psychological biases have actually been popularized as fallacies: notable examples include the so called conjunction fallacy (the tendency to consider the probability of each conjunct as lower than the probability of their conjunction; Tversky & Kahneman, 1982) and the gambler’s fallacy (the tendency to think that past random events influence the probability of future random events; Tversky & Kahneman, 1974). To avoid unnecessary confusion, I propose we reserve the label “fallacies” to the kind of argumentative phenomena analyzed in this paper, while reserving the name “biases” for their more recent psychological cousins. The two phenomena are related but far from identical.

certainly there is no proof of their automaticity—in fact, upon reflection the only clear parallel in the two notions is between the inflexibility of biases and the alleged incorrigibility of fallacies. Thus treating fallacies as argumentative biases would require a drastic (and perhaps interesting) revision of current fallacy theory. While the effort may well bear useful fruits, it would also end up making the notion of fallacy redundant, precisely because it would be supplanted by the more flexible and value-neutral concept of bias. Thus, once again, our analysis suggests that fallacy, as a theoretical construct, may be due for retirement, at long last.

4. Conclusions

The fact that many arguments structurally identifiable as fallacies on the standard approach are not fallacious at all is nowadays firmly established in argumentation theory, and a variety of approaches have endeavored to spell out what demarcate a truly fallacious argument from its structurally similar but ultimately legitimate cousins (for a brief discussion, see section 2). This puts the notion of fallacy in a bit of a quandary, since its denotation is unambiguously pejorative, yet it ends up being used, more often than not, to talk about acceptable argument—hence the bizarre expression “non-fallacious fallacies”. Should the problem be merely terminological, a regimentation would still be in order, but perhaps not terribly urgent. What makes the matter important, however, are the practical implications of the continued use of the concept of fallacy, especially in education.

Indeed, fallacies still loom large in textbooks, both of logic (as a case in point, take the classic Copi, Cohen, & McMahon, 2010, now in its 14th edition) and of critical thinking (a recent authoritative exemplar is Tindale, 2007). While the latter typically offer a much more nuanced analysis of fallacies, along the lines discussed in this paper, the former popularize among students an outdated and misleading treatment of the subject matter. By way of example, consider the following argument that Copi et al. (2010) propose as an example of fallacious DA:

If Carl embezzled the college funds, then Carl is guilty of a felony. Carl did not embezzle the college funds. Therefore Carl is not guilty of a felony.

Only someone with logical blinders on [an expression I gladly borrow from Cohen, (2013), where he was commenting on another bad example from the same textbook] could fail to see that this argument, although obviously deductively invalid, also embodies a perfectly reasonable pattern of inference that may very well be appropriate in a variety of contexts. In fact, in the most obvious context of application, i.e., legal proceedings, this is the *only* acceptable conclusion that could be drawn from the evidence, given that in most legal systems the presumption of innocence places the burden of proof on the prosecutor. This is especially pertinent since “being guilty of a felony” is, on the face of it, a legal concept, not just a mere statement of facts. In order to be guilty of a felony, Carl has not only to have committed the relevant deed, but also to have been proven responsible for it beyond reasonable doubt—which is precisely what is not achieved according to the argument’s premises, therefore fully justifying the conclusion that Carl is not guilty of a felony.

This is not just nitpicking and taking cheap shots at a well-respected textbook. General concerns about the usefulness of fallacy theory to promote critical thinking education have been voiced by scholars who are fully committed to improve, rather than stifle, such educational aim

(e.g., Hitchcock, 1995; Feldman, 2009).¹³ The main reason behind these worries was empirical, since fallacy theory was noted to have little measurable impact on critical thinking proficiency. Now the argumentative theory of reasoning proposed by Mercier and collaborators justifies further concerns: insofar as the main negative side-effect of our cognitive inclinations is to make us exclude too much social information (since the communicative arms race makes people overly skeptical of each other's claims and agendas), priming us to be even more skeptical of arguments may actually be deleterious, rather than just useless—a provocative empirical hypothesis, which should be tested in future studies.

Thus teaching students that something like the embezzling argument is a fallacy does a severe disservice to critical thinking education. Not only because it trivializes and misrepresents the sophisticated understanding of fallacies characteristic of contemporary argumentation theory, but also because it perpetrates the view of fallacies as “silly mistakes that dumb people do”. This is actually intrinsic to the very concept of fallacy, and gives yet another reason to discontinue its usage: a fallacy is something bad we do to ourselves, a pit we fall in of our own free will, a mistake of which we are responsible. As discussed, the alleged frequency with which this happens has been vastly exaggerated: as Woods (2013) put it, “if we attended to the empirical record of the reasoning behavior of individual agents, it would become quickly apparent that, with the [standard list of fallacies] as our guide, beings like us hardly ever commit *them*” [emphasis in original] (p. 139).

So why keep insisting that we are so desperately prone to error? Most likely, to contain the unavoidable hubris that would overcome us if we were to realize that, all things considered, we are not that bad at reasoning, especially when we are free to reason in the appropriate social context. I happen to agree that such hubris would be dangerous and ultimately misguided: being on average decent reasoners does not make us infallible (far from it, in fact), so we must ensure people have a good grasp of their own limitations, as well as their own potential. But the fallibility caveat should be presented in a very different light, based on the results of our discussion. Fallacies, if we insist to keep teaching about them, should no longer be framed as attractive mistakes that dumb reasoners are prone to make, but rather as *points of vulnerability that sophisticated arguers may try to exploit* to their advantage, and possibly to our chagrin. The source of such vulnerability is, interesting enough, our very ability to discriminate between real fallacies and alleged ones. As Godden and Zenker (2015) correctly pointed out, that discrimination hinges on assumptions that are external to the textual material, and thus more malleable to strategic manipulation.

This perhaps offers a way of vindicating their reluctance to abandon the concept of fallacy, but it also requires a radical reinterpretation of their claim: the point is not that the gappiness of fallacies is especially problematic, but rather that, as any other kind of gappiness, it opens the way to the smart exploitation of our interpretative practices. The same holds, once again, for enthymemes in general, which indeed are known to be amenable to strategic tampering— advertising campaigns and political speeches offer a wealth of excellent examples. This is neither a reason to stop using enthymemes, nor to look at fallacy discrimination with special concern. But it does justify rethinking how we teach fallacies to students: instead of hammering them about mistakes that most likely they were not committing in the first place, we should train them to exert due diligence in assessing the arguments presented by their fellows,

¹³ Yet another critique of the traditional pedagogy of fallacies was proposed by Hundleby (2010), but from a very different angle—namely, a concern for the exceedingly adversarial and authoritarian nature of that pedagogy.

with an emphasis on how to spot and neutralize strategic derangements of enthymematic reasoning.

This will require a shift in how we conceive fallacies: from the proneness-to-error approach to the vulnerability-to-manipulation view. It may seem a small shift, but it is not, especially in terms of the pedagogical methods used to teach about fallacies—or whatever we end up calling them. Luckily, some useful suggestions in this direction are to be found in the literature: e.g., Johnson and Blair’s (1977) well-established conception of critical thinking as “logical self-defense”, and the more recent notion of epistemic vigilance (Sperber et al., 2010). This, I suggest, is the direction that fallacy theory ought to explore—in fact, possibly the *only* direction that might justify its continued success, hopefully in a new, improved incarnation.

References

- Boudry, M., Pigliucci, M., & Paglieri, F. (2015). The fake, the flimsy, and the fallacious: Demarcating arguments in real life. *Argumentation* 29 (4), 431-456.
- Brinton, A. (1995). The *ad hominem*. In: H. V. Hansen & R. C. Pinto (Eds.), *Fallacies: Classical and Contemporary Readings* (pp. 213-222). University Park: Penn State University Press.
- Cohen, D. (2013). Virtue, in context. *Informal Logic* 33 (4), 471-485.
- Copi, I., Cohen, C., & McMahon, K. (2010). *Introduction to Logic* (14th ed.). New York/London: Routledge.
- Eemeren, F. van, & Grootendorst, R. (1982). Unexpressed premisses: Part I. *Journal of the American Forensic Association* 19 (2), 97-106.
- Eemeren, F. van, & Grootendorst, R. (1983). Unexpressed premisses: Part II. *Journal of the American Forensic Association* 19 (4), 215-225.
- Eemeren, F. van, & Grootendorst, R. (1995). The pragma-dialectical approach to fallacies. In: H. V. Hansen & R. C. Pinto (Eds.), *Fallacies: Classical and Contemporary Readings* (pp. 130-144). University Park: Penn State University Press.
- Ennis, R. (1982). Identifying implicit assumptions. *Synthese* 51, 61-86.
- Evans, J. St. B. T. (1998). Matching bias in conditional reasoning: Do we understand it after 25 years? *Thinking and Reasoning* 4, 45-82.
- Feldman, R. (2009). Thinking, reasoning, and education. In: H. Siegel (Ed.), *The Oxford Handbook of Philosophy of Education* (pp. 67-82). Oxford: Oxford University Press.
- Finocchiaro, M. (1981). Fallacies and the evaluation of reasoning. *American Philosophical Quarterly* 18 (1), 13-22.
- Floridi, L. (2009). Logical fallacies as informational shortcuts. *Synthese* 167, 317-325.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives* 19 (4), 25-42.
- Geis, M., & Zwicky, A. (1971). On invited inferences. *Linguistic Inquiry* 2, 561-566.
- Gerritsen, S. (2001). Unexpressed premisses. In: F. H. van Eemeren (Ed.), *Crucial Concepts in Argumentation Theory* (pp. 51-79). Amsterdam: Sic Sat.
- Gigerenzer, G., & Selten, R. (Eds.). (2001). *Bounded Rationality: The Adaptive Toolbox*. Cambridge: The MIT Press.
- Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The Foundations of Adaptive Behavior*. New York: Oxford University Press.
- Godden, D., & Walton, D. (2004). Denying the antecedent as a legitimate argumentative strategy: A dialectical model. *Informal Logic* 24, 219-243.

- Godden, D., & Zenker, F. (2015). Denying antecedents and affirming consequents: The state of the art. *Informal Logic* 35 (1), 88-134.
- Goodwin, J. (1998). Forms of authority and the real *ad verecundiam*. *Argumentation* 12 (2), 267-280.
- Govier, T. (1987). *Problems in Argument Analysis and Evaluation*. Dordrecht: Foris.
- Hahn, U., & Oaksford, M. (2006). A Bayesian approach to informal argument fallacies. *Synthese* 152, 207-236.
- Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review* 114, 704-732.
- Hahn, U., Harris, A. J. L., & Corner, A. (2009). Argument content and argument source: An exploration. *Informal Logic* 29 (4), 337-367.
- Hahn, U., Oaksford, M., & Harris, A. J. L. (2012). Testimony and argument: A Bayesian perspective. In: F. Zenker (Ed.), *Bayesian Argumentation* (pp. 15-38). Dordrecht: Springer.
- Hahn, U., Oaksford, M., & Harris, A. J. L. (2013). Rational inference, rational argument. *Argument & Computation* 4, 21-35.
- Hamblin, C. (1970). *Fallacies*. London: Methuen.
- Hansen, H. V. (2015). Fallacies. *Stanford Encyclopedia of Philosophy*. Retrieved May 29, 2015 from <http://plato.stanford.edu/entries/fallacies/>
- Harris, A. J. L., Hsu, A. S., & Madsen, J. K. (2012). Because Hitler did it! Quantitative tests of Bayesian argumentation using *ad hominem*. *Thinking & Reasoning* 18 (3), 311-343.
- Hintikka, J. (1987). The fallacy of fallacies. *Argumentation* 1 (3), 211-238.
- Hitchcock, D. (1985). Enthymematic arguments. *Informal Logic* 7 (2-3), 83-97.
- Hitchcock, D. (1995). Do the fallacies have a place in the teaching of reasoning skills or critical thinking? In: H. V. Hansen & R. C. Pinto (Eds.), *Fallacies: Classical and Contemporary Readings* (pp. 319-327). University Park: Penn State University Press.
- Hitchcock, D. (1998). Does the traditional treatment of enthymemes rest on a mistake? *Argumentation* 12, 15-37.
- Hitchcock, D. (2006). Informal logic and the concept of argument. In: D. Jacquette (Ed.), *Philosophy of logic. Handbook of the Philosophy of Science* (Vol. 5, pp. 101-129). Amsterdam: Elsevier.
- Hitchcock, D. (2007). Why there is no *argumentum ad hominem* fallacy. In: F. H. van Eemeren & B. Garssen (Eds.), *Proceedings of the Sixth Conference of the International Society for the Study of Argumentation* (Vol. 1), June 2006, (pp. 615-620). Amsterdam: Sic Sat.
- Hundleby, C. (2010). The authority of the fallacies approach to argument evaluation. *Informal Logic* 30 (3), 279-308.
- Johnson, R., & Blair, A. (1977). *Logical Self-Defense*. Toronto: McGraw-Hill Ryerson.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Korb, K. (2004). Bayesian informal logic and fallacy. *Informal Logic* 24, 41-70.
- Kuhn, D. (1991). *The Skills of Arguments*. Cambridge: Cambridge University Press.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin* 108 (3), 480-498.
- Laughlin, P., & Ellis, A. (1986). Demonstrability and social combination processes on mathematical intellectual tasks. *Journal of Experimental Social Psychology* 22, 177-189.
- Levi, D. S. (1999). The fallacy of treating the *ad baculum* as a fallacy. *Informal Logic* 19 (2-3), 145-159.
- Mackenzie, P. T. (1980). *Ad hominem* and *ad verecundiam*. *Informal Logic* 3 (3), 9-11.

- Massey, G. (1981). The fallacy behind fallacies. *Midwest Studies in Philosophy* 6 (1), 489-500.
- Mercier, H. (2010). The social origins of folk epistemology. *Review of Philosophy and Psychology* 1 (4), 499-514.
- Mercier, H. (2013). Our pigheaded core: How we became smarter to be influenced by other people. In: K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and Its Evolution* (pp. 373-398). Cambridge: MIT Press.
- Mercier, H., & Sperber, D. (2009). Intuitive and reflective inferences. In: J. St. B. T. Evans & K. Frankish (Eds.), *In Two Minds: Dual Processes and Beyond* (pp. 149-170). New York: Oxford University Press.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences* 34 (2), 57-74.
- Mercier, H., Trouche, E., Yama, H., Heintz, C., & Girotto, V. (2015). Experts and laymen grossly underestimate the benefits of argumentation for reasoning. *Thinking & Reasoning* 21 (3), 341-355.
- Mizrahi, M. (2010). Take my advice—I am not following it: *Ad hominem* arguments as legitimate rebuttals to appeals to authority. *Informal Logic* 30 (4), 435-456.
- Moldovan, A. (2009). Pragmatic considerations in the interpretation of denying the antecedent. *Informal Logic* 29, 309-326.
- Moshman, D., & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking & Reasoning* 4 (3), 231-248.
- Nickerson, R. (1998). Confirmation bias: A ubiquitous phenomena in many guises. *Review of General Psychology* 2, 175-220.
- Oaksford, M., & Hahn, U. (2004). A Bayesian approach to the argument from ignorance. *Canadian Journal of Experimental Psychology* 58, 75-85.
- Paglieri, F. (2014). Accepted by whom? On the empirical roots of Aristotle's dialectic. *Revue Internationale de Philosophie* 270, 393-402.
- Paglieri F., & Woods J. (2011). Enthymematic parsimony. *Synthese* 178, 461-501.
- Perkins, D., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In: J. Voss, D. Perkins & J. Segal (Eds.), *Informal Reasoning and Education* (pp. 83-105). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pohl, R. (Ed.). (2004). *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement, and Memory*. Hove/New York: Psychology Press.
- Rahwan, I., Krasnoshtan, D., Shariff, A., & Bonnefon, J-F. (2014). Analytical reasoning task reveals limits of social learning in networks. *Journal of the Royal Society Interface* 11 (93), 20131211.
- Scriven, M. (1987). Fallacies of statistical substitution. *Argumentation* 1, 333-349.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind and Language* 25 (4), 359-393.
- Stalnaker, R. (1968). A theory of conditionals. *Studies in Logical Theory: American Philosophical Quarterly Monograph* 2, 98-112. Oxford: Blackwell.
- Stanovich, K., & West, R. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning* 13 (3), 225-247.
- Stone, M. (2012). Denying the antecedent: Its effective use in argumentation. *Informal Logic* 32, 327-356.
- Tindale, C. (2007). *Fallacies and Argument Appraisal*. Cambridge: Cambridge University Press.
- Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

- Trouche, E., Johansson, P., Hall, L., & Mercier, H. (forthcoming). The selective laziness of reasoning. *Cognitive Science*, doi:10.1111/cogs.12303
- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General* 143 (5), 1958-1971.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science* 185(4157), 1124-1131.
- Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In: D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 153-160). Cambridge: Cambridge University Press.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90 (4), 293-315.
- Walton, D. (1995). *A Pragmatic Theory of Fallacy*. Tuscaloosa: University of Alabama Press.
- Walton, D. (1997). *Appeal to Expert Opinion: Arguments from Authority*. University Park: Penn State University Press.
- Walton, D. (1998). *Ad Hominem Arguments*. Tuscaloosa: University of Alabama Press.
- Walton, D. (1999). The appeal to ignorance, or *argumentum ad ignorantiam*. *Argumentation* 13 (4), 367-377.
- Walton, D. (2000). *Scare Tactics: Arguments that Appeal to Fear and Threats*. Dordrecht: Kluwer.
- Walton, D. (2001). Enthymemes, common knowledge, and plausible inference. *Philosophy and Rhetoric* 34 (2), 93-112.
- Walton, D. (2008). The three bases for the enthymeme: A dialogical theory. *Journal of Applied Logic* 6 (3), 361-379.
- Walton, D., & Krabbe, E. (1995). *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. Albany: SUNY Press.
- Wason, P. C. (1966). Reasoning. In: B. Foss (Ed.), *New Horizons in Psychology* (Vol. I, pp. 106-137). Harmondsworth: Penguin.
- Woods, J. (1998). *Argumentum ad baculum*. *Argumentation* 12 (4), 493-504.
- Woods, J. (2013). *Errors of Reasoning. Naturalizing the Logic of Inference*. London: College Publications.
- Zenker, F. (Ed.). (2013). *Bayesian Argumentation: The Practical Side of Probability*. Berlin: Springer.