

---

# Book Review

Review of *Errors of Reasoning. Naturalizing the Logic of Inference*, by John Woods, Studies in Logic 45, College Publications, London, 2013, ISBN 978-1-84890-114-8.

John Woods' *Errors of Reasoning* (EoR) flies in the face of logical orthodoxy in an astonishing number of ways. Its 550 pages make numerous claims on a great many subjects. Due to space limitations, this review will focus only on some of them. I nevertheless hope to give the general flavour of this thought-provoking book.

The official core topic of EoR is the one of logical fallacies. And the core claim is a critical one: the traditional theory of fallacies is radically wrong. But the background for this bold negative statement is a sophisticated and highly revisionary philosophical view of human reasoning. Such a view, it seems to me, is the most interesting aspect of the book, its critique of traditional fallacy theory being, in a sense, but its main local application. I will begin by explaining the gist of Woods' critical claim on fallacies. I will then explore the background approach inspiring Woods' stance, which occupies, roughly, EoR's first half. Finally, I will give an example of the strategies used by Woods to substantiate the critical claim via the background view.

The heart of Woods' critique lies in what he calls the *concept-list misalignment thesis* (p. 6). He identifies a mainstream concept of logical fallacy and a list of items traditionally labelled as fallacies. He then shows via a detailed case-by-case investigation how most of the items in the list do not belong in the extension of the concept.

As for the targeted concept, it is called the EAUI (read it as 'Yowee': p. 135) concept of logical fallacy: fallacies are Errors of reasoning, which are Attractive to reasoners (in that they easily fall prey of them), Universal (i.e., extremely common) and Incurable (even after the given fallacy's fallacious nature has been explained to the reasoner, s/he is still prone to commit it). As for the list considered in EoR, it includes traditional deductive and inductive fallacies like *ad baculum*, *ad hominem*, *ad populum*, *ad ignorantiam*, *ad verecundiam*, affirming the consequent, denying the antecedent, hasty generalization, equivocation or *quaternio terminorum*, the gambler's fallacy and *post hoc ergo propter hoc*. The EAUI conception gives a conjunctive characterization, so Woods has a disjunctive task: of each targeted (would-be-)fallacy, show that, contrary to tradition, it is either not attractive, or not universal, or not incurable—or, *not an error*. This last bit takes us to Woods' second, possibly even more interesting claim on fallacies after concept-list misalignment: EoR's *cognitive virtue thesis* (p. 7) has it that most of the traditional fallacies should not be labelled as errors, but are, on the contrary, virtuous ways of reasoning. Now one expects what Lewis called 'incredulous stares': how can denying the antecedent or *ad ignorantiam* not be logical mistakes? There is no way for 'If *A* then *B*' and not-*A* to entail not-*B*, or for 'There's no evidence that not-*A*' to entail *A*, in any logically decent sense of 'entail'. But according to Woods, much mainstream logic has got the standards of logical decency wrong: he calls, in fact, for a rethinking of what counts as error in human reasoning.

The alternative view of reasoning and its errors proposed by Woods to substantiate his concept-list misalignment and cognitive virtue theses is based on a broad and ambitious naturalization project, introduced in Chapter I and expanded to specific topics mostly in Chapters II to V. This is inspired by a number of traditions neglected or left under-developed within the mainstream 20th Century logic, ranging from naturalized epistemology to cognitive science and to some

non-monotonic and abductive approaches to human reasoning. The new logic is to be agent-based in a radical sense. Approaches to doxastic logic, to belief representation and revision, etc., already include reference to (cooperative, finite, fallible) agents. But Woods wants to make room for agents freed from *any* kind of logical idealization: a view which ‘pays attention to what people are like, how they are put together and what they get up to when they reason’ (p. 11). Real-world cognitive agents like us are not just resource-bound and fallible, but also driven by a cognitive agenda of practical and contingent interests. To take them seriously, we need a logic of reasoning that accepts serious input not only from epistemology, but also from psychology. We need to make a lot of the difference (pp. 24–26) between consequence-having (in the various senses of logical consequence, which ‘occurs in logical space’) and consequence-drawing (which occurs ‘in a reasoner’s mind’). It may turn out that the consequences a group of premises has often fail to be the consequences the relevant cognitive agent *ought to* draw for its own purposes. Even in the realm of consequence-having, most of human reasoning is what Woods calls ‘third-way reasoning’: a blend of non-monotonic, default, *ceteris paribus*, agenda-relevant, inconsistency-adaptive reasoning, for which neither standard deductive validity nor inductive strength are appropriate canons of assessment.

Chapter II is largely concerned with methodological matters, drawing a broad picture of how the naturalistically inclined theory-builder in the business of putting this new logic of reasoning together should select and respect the relevant data, and of what should count as the theory’s actually explaining them. Woods here stresses how, in an agent-centred theory of reasoning, an unavoidable trade-off is to be expected between respect for empirical data on the one hand, and the need for idealization, formalization and model-building on the other. The main claim of the Chapter is to the effect that the naturalized logic we are after should have the normal and the normative converge: ‘it is typically the case that how we reason from premises to conclusions in real life circumstances is either *accurate* or *apt* or both’ (p. 54). A central issue for naturalized views in most areas of philosophy is how to reconcile their descriptive make-up with our normative intuitions. Woods’ general strategy to defend the cognitive virtue thesis roots our inferential practices in our natural cognitive make-up: our mostly doing what we ought to do in reasoning is a case of our mostly succeeding in knowing what we ought to know. This follows from our being cognitive agents which underwent natural selection: our cognitive devices’ working properly in most circumstances is a *conditio sine qua non* of our survival and prosperity.

A naturalized logic of reasoning is then rooted in a naturalized epistemology: this is presented in Chapters III to V. Many points made by Woods draw on views in informal naturalized epistemology, from Goldman’s reliabilism to the works of Kornblith; but he puts this material to new service. Woods’ favourite epistemological view, called the Causal Response model (CR, p. 93), is presented by contrasting it with a more traditional philosophical conception of knowledge, labelled the Command and Control reasoning model (CC, p. 99). This is the conception in Plato’s tradition: knowledge is justified true belief, with the justification bit complicated in order to reply to Gettier cases (pp. 111–115). The CC model takes justification as a conscious procedure, where the agent knowing that *A* is expected to be able to actively make a case on behalf of such knowledge when facing sceptical scenarios. Woods takes the CC model as forcing the naturalistically unacceptable conclusion that we do not know much of anything, with a stark divergence between the demands of normativity and those of reality (pp. 90–91 and 98–100). So much the worse, as Woods’ *modus tollens* goes, for the very importance of the justification condition for knowledge. The CR model has it that knowledge is true belief rooted in the correct functioning of our cognitive devices. Then, in compliance with the convergence of natural and normative, ‘human beings have knowledge, lots of it’ (p. 86). The CR model also allows most of our knowledge to be ‘dark knowledge’ (p. 124): true

beliefs produced in automatic mode precisely by our cognitive devices' being causally triggered and firing in the appropriate ways. Woods proposes to extend the picture to belief-update, reasoning in general, and third-way reasoning in particular: this is in most cases left to our spontaneous devices, 'unconscious, sublinguistic, inattentive, involuntary, automatic, effortless, non-semantic, computationally luxuriant, parallel, and deep' (p. 123). The mainstream AGM theory of belief-change is then taken into account, and also found guilty of logical idealization insofar as AGM belief revision is driven by classical logic and closed under classical consequence (pp. 152–153). An alternative explanation of how we revise our beliefs in the light of (inferences from) new information 'rests with the natural sciences of cognition. It is an answer for the lab to produce [...]. One of the reasons to like the CR-model is that it is so at home with the idea of how little our adeptness at knowing things owes to a stateable awareness of how to do it' (p. 183). Finally, most of our cognitive mistakes in the CR approach are due to misinformation rather than to bad reasoning (Chapter VI, esp. pp. 188–192).

With this background settled, Woods devotes most of the remaining Chapters of EoR to a case-by-case examination of various (would-be-)fallacies. Many of the traditional fallacies *are* deductively invalid or inductively weak, of course. But we now have the philosophical bases for a conception of third-way reasoning which makes it inappropriate to assess the fallacies by means of standards of deductive validity or inductive strength. Chapters VII and VIII describe the workings of defeasible non-monotonic logics in detail, highlighting the epistemically loaded character of their consequence relations (pp. 235ff) and the strongly irregular features of the notion of defeasibility vis-à-vis other better behaved concepts, such as probability and (normal) modalities like necessity and possibility (pp. 271ff). While Woods acknowledges the merits of various traditions in non-monotonic logic, his main conclusion is still negative: we should, in an important sense, let the very notion of *consequence* go. His 'downgrading consequence' conclusion is that 'when a third-way premiss-conclusion reasoning  $R$  is nonmonotonic, the construal of  $R$  as consequence makes no load-bearing contribution to the logic of  $R$ ' (p. 281). Conclusion-drawing is a more complex six-place relation, intrinsically *causal* (contrast *propositional*),  $\langle X, I, \Delta, A, \alpha, D \rangle$ , including a cognitive agent  $X$ , the information  $I$  the agent reasons on, a background database  $\Delta$  of information available to the agent, a cognitive agenda  $A$  of duties for the agent, a conclusion  $\alpha$ , obtained after applying the background  $\Delta$  to  $I$ , and a disposition  $D$  to answer to justification requests for  $\alpha$ 's being drawn (pp. 285–286). If the various families of non-monotonic and abductive logics make for the most promising strategies on the market to account for third-way reasoning, a significant restructuring of them will be needed if they are to become capable of accounting for conclusion-drawing structural relations of this kind.

While in the business of spelling out all of this, Woods also establishes that denying the antecedent is not to be considered an EAUI fallacy (pp. 252–254). This gives us an example of how he supports the concept-list misalignment thesis and the cognitive virtue thesis at once. Traditionally, denying the antecedent is taken as the mistake of assuming a conditional and the negation of its antecedent, and from this inferring the negation of the consequent. But if we want to make sense of this in the light of our CR epistemology and of the convergence of the normal and the normative, our best strategy for Woods is to *re-interpret* what we actually do in situations that the traditional view reads as committing the fallacy. And what we do is plausibly understood as retraction of defeasibly drawn conclusions. We have been sticking to information in  $\Sigma$ 's entailing  $B$ , but new information contradicting  $\Sigma$  (falsifying items in it, say, some  $A \in \Sigma$ ) now occurs; thus, 'not- $A$ ', where "'not" expresses retraction, not negation' (p. 253). Now it is plausible to make a move to the effect that  $B$  be retracted too: 'not- $B$ '. Then denying the antecedent is 'out of the fallacy business' (p. 254). In the contexts in which it is most likely

to occur, it should be understood as a cognitively healthy retraction move, while it can only be unhealthy in a limited number of circumstances, whereby it misses the fallacy requirement of universality.

A similar strategy is pursued in Chapter XI, to rehabilitate affirming the consequent, traditionally taken as the mistake of having a conditional and its consequent, and from this inferring the antecedent. Woods proposes, again, to properly interpret the supposed fallacy. When this is done, it turns out to be a form of abduction in most of the real-life reasoning contexts in which it occurs. Woods gives here a précis (pp. 368ff) of the sophisticated account of abduction proposed elsewhere by Dov Gabbay and himself (see [1]). If we want to stick to the original spirit of the Peircean notion, according to Woods we have to admit that abduction is often *ignorance-preserving* (p. 378): our postulating a hypothesis,  $H$ , such that, to put it counterfactually, if  $H$  were true then the surprising fact  $F$  would be a matter of course, does not as such remove our ignorance on  $H$ . Rather, it makes of  $H$  a conjecture, put on the table as a target for further investigation. Now suppose we interpret affirming the consequent as having the form: (1) If  $H$  were the case, (actually surprising)  $F$  would (counterfactually) be a matter of course; (2) But lo,  $F$  is the case (the surprising, observed fact); therefore (3) Defeasibly, it is reasonable to conjecture that  $H$ . What was an EAUI fallacy in the traditional view has turned into healthy abduction (pp. 382–384). A similar fate awaits begging the question (Chapter XII), *ad hominem* (Chapter XIII) and the gambler's fallacy (Chapter XIV): none of these counts as a case of EAUI. Each is either not incorrigible; or such that its frequency does not make it satisfy the universality condition; or such that it can be interpreted as a benign form of consequence-drawing for resource-bound, agenda-driven, naturalistically conceived cognitive agents like us.

In the final Chapter XV, Woods discusses some objections to his view and misgivings about the consequences of its being correct. Perhaps the most interesting question taken up here asks whether, after all, the EAUI concept itself is in disarray—which seems plausible, after all of EoR's destructive work. But for Woods not even this is the case: there may well be a non-traditional list of instances for EAUI. An example (pp. 513–514) is Powers' Paradox: this concludes for the (natural language: 'if and only if') equivalence between the material conditional, ' $A \supset B$ ', and the natural language conditional, 'If  $A$ , then  $B$ ', by assuming that if the former is true, then the latter is (this last assumption being natural language 'if ..., then ...' again). Powers' Paradox, says Woods, is a mistake, for the material conditional is too bad to correspond to the natural language 'if ..., then ...', due to the paradoxes of relevance. It is attractive, though, because the mistaken assumption 'If " $A \supset B$ ", then if  $A$ , then  $B$ ' is easily confused with material *modus ponens*, 'If " $A \supset B$ " and  $A$ , then  $B$ '. It is also widespread, and such that people tend to fall into it again and again: logic students assent to Powers' Paradox when taking logic classes from all over the world. This hints at a positive post-EoR research project: find out the real EAUI fallacies according to the newly proposed standards of third-way reasoning.

A few final words on EoR's general significance. As we have seen, this is an essentially philosophical and largely critical book. Hence its main limitation—which Woods explicitly declares. EoR gives us a general methodological framework for the new, naturalized logic of reasoning, and a large amount of reasons for thinking that the traditional view on fallacies is deeply wrong. We should not expect to find in the book, though, a presentation of the technical details of the logic of third-way reasoning. As Woods makes clear from the start, the book puts emphasis 'more on the to-be-done than on the done', and is 'more a prolegomenon to a naturalized logic of error than the finished product' (p. 41). Sympathetic mainstream logicians may then ask the follow-up question: 'Ok, now what?', barely hiding their persistent need for the mathematically precise modelling of whatever they consider worthy of attention. The good news for such logicians is that Woods' long-time logical

teammate, Dov Gabbay, is working on the companion volume of EoR, whose provisional title, *Formalizing the Logic of Error*, is likely to trigger their interest. Reading EoR has triggered mine for sure.

FRANCESCO BERTO  
University of Aberdeen  
E-mail: [f.berto@abdn.ac.uk](mailto:f.berto@abdn.ac.uk)

## Reference

- [1] D. M. Gabbay and J. Woods. *The Reach of Abduction: Insight and Trial*. North Holland, 2005.