

Chapter 4

Give the Standard Treatment of Fallacies a Chance! Cognitive and Rhetorical Insights into Fallacy Processing



Steve Oswald and Thierry Herman

4.1 Introduction

The identification of normative standards of argument quality can be said to be one of the cornerstones of argumentation scholarship, so much so that “[t]he degree to which a theory of argumentation makes it possible to give an adequate treatment of the fallacies can even be considered as a litmus test of the quality of the theory” (van Eemeren et al., 2014, p. 25). Across the ages, various normative accounts have come up with different standards meant to discriminate good from bad arguments. In his influential work on fallacies and on the history of their study, Hamblin (1970) referred to most accounts of fallacious argumentation available in his time as “the standard treatment”. Under this account, a fallacious argument is an argument that “seems to be valid but is not so” (Hamblin, 1970, p. 12). Hamblin, and others in his wake, abundantly criticised this definition, notably for its inability to identify stable and accurate norms for what should count as (un)acceptable argumentation.

While we acknowledge the inadequacy of the standard treatment when it comes to assessing the normative dimension of argumentation, we suggest that its definition of fallacious arguments sheds light on another issue of interest to argumentation scholars, namely rhetorical effectiveness.¹ In this chapter, we thus presuppose the

¹By rhetorical effectiveness, we mean any type of influence that the argument might have on the addressee and that is consistent with the speaker’s persuasive intentions.

S. Oswald (✉)

English Department, University of Fribourg, Fribourg, Switzerland
e-mail: steve.oswald@unifr.ch

T. Herman

Institute of Information and Language Science,
University of Neuchâtel, Neuchâtel, Switzerland
e-mail: thierry.herman@unine.ch

School of French as a Foreign Language, University of Lausanne, Lausanne, Switzerland

© Springer Nature Switzerland AG 2020

F. H. van Eemeren and B. Garssen (eds.), *From Argument Schemes to Argumentative Relations in the Wild*, Argumentation Library 35,
https://doi.org/10.1007/978-3-030-28367-4_4

existence of normative standards to assess argument quality, but we wish to move away from normative concerns in order to explore the question of rhetorical effectiveness. Specifically, we discuss the fundamental intuition regarding the way fallacies are processed, and which is contained in Hamblin's definition: if fallacies are misidentified as valid (or acceptable/reasonable/unproblematic), then it means that their fallaciousness remains undetected. This, in itself, justifies an inquiry into how they manage to remain undetected. We propose to outline the contours of this inquiry by drawing on an explanatorily adequate cognitive pragmatic account of meaning.

In doing so, we will defend the claim that rhetorically effective fallacies cognitively constrain the addressee's argumentative evaluation of the message, so as to bypass or withstand critical testing. This involves, on behalf of the speaker, the formulation of argumentative messages that are likely to strengthen the content of the fallacy to make it epistemically stronger and cognitively easy to process and/or weaken potential or actual refutations on the same two dimensions (epistemic strength and ease of processing). In a nutshell, this contribution (i) develops the rationale for envisaging rhetorical effectiveness in terms of foregrounding and backgrounding cognitive processes, and (ii) explains how the linguistic choices made in fallacious arguments are likely to decisively constrain these cognitive processes in their attempt to secure rhetorical effectiveness.

Section 4.2 starts with a discussion of the standard treatment of fallacies and reports on some normative issues it has faced, before proceeding to defend a rhetorical reading of the definition of fallacy it offers in an attempt to salvage it. Section 4.3 offers a cognitive and reductionist account of how argumentative evaluation unfolds, with specific focus on fallacies. It then outlines an account of what rhetorical effectiveness is, from a cognitive perspective, focusing on parameters which determine informational selection. Section 4.4 illustrates how these mechanisms play out by discussing various backgrounding and foregrounding strategies at play in a courtroom dialogue taken from the 2009 film *Public Enemies*, directed by Michael Mann.

4.2 The Standard Treatment of Fallacies: Problems and Opportunities

4.2.1 Normative Issues

Hamblin, in his ground-breaking study on fallacies (Hamblin, 1970) observes that most available accounts of fallacious argumentation in the logic textbooks of his time seem to have much in common. The vast majority, according to him, are very traditional (few of them go beyond Aristotle's typology) and they all present similar typologies of finite sets of arguments, which are said to be fallacious. He dubs these accounts 'the standard treatment of fallacies' and levels devastating criticism against it, as, according to him, it is highly unsatisfactory and problematic on a number of counts. To quote his own words, this treatment is "as debased, worn-out and

dogmatic a treatment as could be imagined—incredibly tradition-bound, yet lacking in logic and historical sense alike, and almost without connection to anything else in modern logic at all” (Hamblin, 1970, p. 12). Accordingly, he considers the landscape of fallacy research to be so gloomy that it has prevented argumentation scholarship from elaborating any serious account of fallacious argumentation.

Hamblin’s criticism begins with the identification of what he takes to be a common definition of fallacy within the standard treatment. Under this account, a fallacious argument is an argument that “*seems to be valid but is not so*” (Hamblin, 1970, p. 12, original emphasis). Since this is a widespread definition, any critique of the definition inevitably disqualifies all existing accounts which can be said to take it on board. Hamblin first notes that, while this definition should make us expect researchers to focus on either what makes fallacious arguments seem valid or on what makes them invalid, virtually no one has risen up to the challenge. Moreover, as Hamblin, but also many others after him, pointed out, adopting such a definition leads to a number of theoretical and methodological problems, among which we find the following:

- the definition only seems to capture a restricted set of formal fallacies and thus fails to capture most types of fallacies (see also van Eemeren, 2001)
- the definition captures arguments which we would be ready to deem fallacious, but which are valid from a logical perspective, such as *petitio principii* (see also van Eemeren, 2001)
- the definition makes fallacy judgements relative (since, in the end, what *seems* valid to someone might not to someone else, see Walton, 2010)
- the definition makes fallaciousness an issue of formal validity, which, again, fails to capture the reality of argumentative fallaciousness (see, again, Walton, 2010, and more generally any account of the rise of informal logic, such as Johnson & Blair, 2006, or Johnson, 2014).

We leave here issues about the inadequacy of the standard treatment in any normative inquiry and refer the reader to the abovementioned literature for an informative discussion of the normative aspects of argument quality. Instead, we wish now to highlight that there are grounds, and in fact argumentatively relevant grounds, to consider that we should not throw out the definition of fallacies given in the standard treatment with the bathwater.

4.2.2 *Salvaging the Standard Treatment: Rhetorical Insights*

We submit here that the definition of fallacy given in the standard treatment contains a basic, yet valuable, intuition about why fallacies might have some appeal to some audiences. In this sense, we propose to depart from logical or dialectical normative considerations on the standard treatment of fallacies and offer a rhetorical reading of the definition. We then draw its implications for the study of argumentation. We propose that Hamblin put his finger on the very mechanism that allows

fallacious arguments to have some rhetorical appeal (which still needs to be defined, see Sect. 4.3.3).

Under the standard treatment, fallacies are defined as invalid arguments that seem valid. That is, that seem valid *to someone*—crucially, to those who fail to identify them as fallacies. It stands to reason that subjects who recognise that, in a line of argumentation offered by a speaker, the conclusion does not follow from the premises brought to bear in the justification would in principle reject it; other things being equal, it is indeed part of one's default epistemic goals not to take on board information that one thinks ill-evidenced.

This first specification allows us to establish that one of the core features of successful fallacies has to do with the fact that there is something about them that makes them undetectable. In other words, successful fallacies are those that manage to prevent their addressees from identifying that, how and/or why they are fallacies. Therefore, a reasonable stance on the issue would be to consider that the rhetorical success of a fallacy is intimately related to the manifestness of its fallacious nature: the less manifest the fallacious status, the more undetectable, and the more rhetorically effective the fallacy. Now, if we are to further explore this line of reasoning, it seems important to be clear on what the notion of manifestness amounts to. Drawing on Sperber & Wilson's relevance theory, we will consider that the manifestness of information is defined in terms of its ability to be represented: "a fact is manifest to an individual at a given time if and only if he is capable of representing it mentally and accepting its representation as true or probably true" (Sperber & Wilson, 1995, p. 39). In the case of fallacious argumentation, it is therefore crucial to its success that any information related to the fallacy's status *qua* fallacy remains non-manifest in whatever information sets the addressee is processing as he encounters the fallacy. We will come back to this as we develop our account below to show how manifestness can be tweaked in the search for rhetorical effectiveness.

Let us now further reflect on the theoretical implications of this idea for the type of inquiry we have in mind. Determining the parameters that make a fallacy undetected ultimately appears to be a psychological issue whose investigation must unfold within some sort of cognitive or psychological framework. This is because what is at stake in that line of inquiry are (i) explanations of why a given set of information fails to be identified as normatively lacking, and, thus, (ii) explanations accounting for why this set of information is not rejected—as it should, in light of its failure to uphold the normative standards in place. In other words, explaining why fallacies can go unnoticed will lead us to wonder about more general cognitive mechanisms underlying the selection (i.e., the inclusion and/or rejection) of information in communicative exchanges.

While the disruptive nature of fallacies is not a new idea in argumentation scholarship, it has not, to our knowledge, been developed so far within a consistent research programme meant to account for their cognitive underpinnings. However, some relevant work paving the way for this investigation deserves to be mentioned here. In her paper on the relationship between fallacies and heuristics, Jackson (1996) notes that fallacies "are not incorrect argument schemes or correct argument schemes applied

incorrectly, but failed diagnostic strategies” (1996, p. 111) and that they are persuasive “because audiences use them as shortcuts to avoid careful thinking about issues” (1996, p. 104). Here Jackson (rightfully) sees the relationship between fallacies, taken as discursive argumentative objects, and cognition as a disruptive one. At the end of the day, fallacies function by perturbing critical thinking processes in order to generate some kind of error—usually consisting in the non-identification of the fallacy as one, and consequently leading to a failure to reject it, this resulting in a so-called “failed diagnostic strategy”. Willard (1995) holds a similar and compatible view when he states that fallacies “distract the persuadee from reflective thinking”, “divert attention from argumentatively weak messages” and “distract attention from matters the analyst takes to be relevant” (1995, pp. 148–149). These considerations resonate with the definition of fallacy offered in the standard treatment in that they seem to specify, albeit without entering into the details, what effect (successful) fallacies have on people’s minds. Essentially, they disrupt evaluative processes and cause their victims to fail to identify them as fallacies, and thus to fail to reject either their conclusion or the justificatory relationship that is offered to support it. In a related area of inquiry, Correia (Correia, 2011, 2014) develops the link between cognitive biases, motivated reasoning and fallacies—albeit from the perspective of the speaker—in order to show that many fallacies are likely to be caused by cognitive biases. We suggest that it is possible to conduct similar reflections about the reception end of the communicative process, namely the addressee, and that one of the fundamental features of fallacies is their ability to exploit the inherent fallibility of cognitive mechanisms.

Let us take stock of all this. What we propose is to tackle the rhetorical effectiveness of fallacious arguments via a cognitive framework, based on the assumption that the study of rhetorical effectiveness ideally lends itself to a cognitive investigation. Moreover, we contend that the definition of fallacy given in the standard treatment contains a valuable starting point for this project: the *appearance condition* (as discussed by Hansen, 2015) stipulates that the condition under which fallacies are successful—whether that is good or bad is a separate question—has to do with their ability to go unnoticed. Thus, what we should explain in order to tackle rhetorical effectiveness is how and why successful fallacies remain undetected (that is less manifest or not manifest at all).

4.3 Why and How Fallacies Are Rhetorically Successful

4.3.1 Cognitive Underpinnings of Argument Evaluation

From a cognitive perspective, and all else being equal, the evaluation of an argument can be modelled as a process in which two sets of information interact: on the one hand, the content of the speaker’s utterance, which contains the premise(s) and the conclusion, and on the other, the set of contextual information against which the

speaker's utterance is processed. In principle, this interaction can yield different outputs in the evaluation process, ranging from complete rejection of the claim to complete acceptance. This characterisation affords two different possible scenarios. In the first scenario, in which the set of contextual information that is mobilised to evaluate the claim and the arguments turns out to be somehow stronger (a notion we characterise more precisely below), the speaker's standpoint is rejected. Conversely, in the second scenario, in which the set of contextual assumptions turns out to be weaker, the standpoint is likely to prevail and to enter the addressee's cognitive environment.²

Let us further specify this picture, starting with a terminological discussion about the nature of information sets that interact in argumentative evaluation. In the evaluative stage of argumentation, where participants try to test the argumentation put forward by the speaker to figure out whether the conclusion should be accepted based on the available and supplied evidence,³ cognitive systems are processing these two information sets against one another: the content of argumentation on one hand and contextual information on the other. We propose to call the set of contextual information brought to bear in argumentative evaluation the *critical context*. The critical context contains information on, among other things, the speaker's (possibly deceptive/manipulative/uncooperative) intentions, potential inconsistencies within the argumentation, potential discrepancies between the argumentation's content and the addressee's beliefs and values, etc. In short, the critical context is likely to contain any piece of relevant information that would help making the fallacious nature of the argumentation manifest.

Let us now try to specify the criteria that are likely to influence argumentative evaluation, which, as previously mentioned, can result in either acceptance or rejection of the speaker's standpoint. Earlier, we noted that acceptance denotes cases in which the content of the speaker's argumentation is stronger than the critical context, and that rejection denotes cases in which the former is weaker than the latter. What exactly do we mean by 'stronger' and 'weaker'? First, we have to note that these two comparative terms are to be understood as related to the selection of information: the stronger a piece of information, the more likely it is to be selected, processed, stored, etc., and, conversely, the weaker the piece of information, the less likely it is to be selected, processed, stored, etc. All else being equal, two main cognitive parameters are responsible for the selection of information in any cognitive process: the first has to do with its reliability, and the second with its processability.

We define the first as *epistemic strength* and the second as *accessibility or availability of critical information*. Epistemic strength predicts acceptance of the conclusion in case the argument turns out to be epistemically superior to critical information

²Sperber and Wilson define the notion of cognitive environment as follows: "A cognitive environment of an individual is a set of facts that are manifest to him" (Sperber & Wilson, 1995, p. 39). We use this characterisation to denote quite loosely the set of information that an individual knows and is capable of bringing to consciousness.

³We note, following Mercier and Sperber (Mercier & Sperber, 2009, 2011, 2017), that evaluation may be intuitive (that is, without consciously attending to the process) or reflective (that is, with reflective awareness of the evaluation).

sets: that is, the content of the argumentation put forward by the speaker is evaluated by the addressee as more likely to be true than whatever critical information he can come up with. We then define the second parameter (processability) as *accessibility/availability of critical information* and predict that failing to access critical information should increase persuasiveness.⁴ In lay terms, and to summarise, there are grounds to consider that an individual will end up accepting the conclusion of an argument when she finds its content particularly plausible in light of the criticism she can think of at that point.

Another way of looking at this characterisation of rhetorical effectiveness is to envisage it as an issue of informational salience, in which persuasive strategies are characterised in terms of their *strengthening* and *weakening* possibilities. *Strengthening* strategies effectively foreground information that is compatible with the content of the claim/argument complex; this can be done by making such information epistemically strong and more accessible. *Weakening* strategies target critical information in order to background it; they achieve this by making it epistemically weaker and less accessible—if not altogether inaccessible.

As argued elsewhere (e.g., Maillat & Oswald, 2009, 2011; Oswald, 2010, 2011, 2014, 2016) these parameters can be linked to Sperber and Wilson's extent conditions of relevance (Sperber & Wilson, 1995, p. 125). While these were originally formulated in relevance theory to specify how addressees arrive at relevant interpretations of speaker meaning through inferential work, we propose that they are equally useful to any account of rhetorical effectiveness (see also Sect. 4.3.3). As per the principle of relevance ["Human cognition tends to be geared to the maximisation of relevance", (Wilson & Sperber, 2002, p. 254)], these extent conditions can indeed be assumed to cover a larger ground and encompass virtually all cognitive activity, including the production and evaluation of argumentation. Under the reductionist perspective we explore here, the evaluation of argumentation is a process in which the cognitive system weighs the epistemic strength of a speaker's argumentation against a (critical) subset of the addressee's cognitive environment—this subset being itself likely to vary in terms of manifestness, depending on its accessibility. In such a framework, we therefore come to define rhetorical effectiveness as a function (i) of the epistemic strength of argumentation and (ii) of the accessibility/availability of critical information. Such a reductionist take on persuasion, we believe, is likely to open new avenues of research if we apply it to the study of fallacies and, more broadly, of argumentation (see Oswald, 2016).

⁴Let us note, in passing, that we assume that this happens on the condition that the content of the argumentation has some minimal degree of plausibility in the addressee's cognitive environment. If the addressee outright rejects the content as being false, or if he simply does not understand it, then accessibility considerations in principle become irrelevant.

4.3.2 *Undetected Fallacies Are Arguments That Are Non-manifest Qua Fallacies*

Within this picture, and assuming that fallacies are normatively flawed arguments the rhetorical success of which rests on the non-manifestness of the flaw, we can hypothesise that fallacies operate specific cognitive constraints on the selection of information in the processes of evaluation. These constraints, in turn, are meant to focus the addressee's attention on specific information sets and to prevent him from mobilising an optimal critical context during evaluation—or to prevent critical evaluation from starting at all. More specifically, we argue that fallacies operate a twofold cognitive constraint on information processing. On the one hand, they aim at strengthening 'fallacy-friendly' information sets, namely those sets of information that support or are compatible with the content of the fallacy. On the other, they aim at obfuscating critical information, or at least at weakening the chances that critical information is brought to bear at the evaluative stage.

Saying that a fallacious argument has been rhetorically effective amounts to saying that either it has *withstood* critical evaluation or that the addressee has accepted it *without* any critical evaluation. In the first scenario, the result is that the content of the argumentation is either epistemically compatible with or epistemically stronger than the critical context against which it is tested. The second scenario can be interpreted in terms of the presence or absence of said critical context. If the fallacy prevails, it could simply be because no critical context is present at the time to defeat it or because the addressee finds no motivation to supply such a critical context; in this case what is at stake is the accessibility of the critical context.

For theoretical and conceptual clarity, we should point out that fallacies are presumably no different from normatively sound arguments when we look at them from a cognitive perspective: both types of arguments are processed by general inferential mechanisms which regulate the selection of information following considerations of relevance, construed in terms of epistemic strength and accessibility. To take van Eemeren's analogy with cats and dogs, fallacies and reasonable or sound arguments are not different animals: instead, fallacies are to be thought of as bad dogs while reasonable or sound arguments are to be thought of as good dogs (van Eemeren, 2015). This analogy also makes sense from a cognitive perspective, as there are no solid reasons (empirical or otherwise) to consider that they rely on different cognitive processes. Furthermore, the unwitting addressee does not identify fallacies as such, meaning that their evaluation unfolds in the standard way—which is, admittedly, also part of the reason they are effective.

However, they do differ in at least two respects, namely (i) their communicative status (in the case of deliberate fallacies, or fallacies by design) and (ii) their foregrounding/backgrounding dynamics. Many researchers have noted that fallacies can be treacherous, misleading and/or deceptive. Aristotle, in his *Sophistical refutations*, noted that fallacies many times take the appearance of valid deductions, and that this is at least partly why they can fool people (see Aristotle, 1955; Hansen, 2015). Ben-
tham (1962) also highlighted the deceptive nature of fallacies, as they cause people

to have false beliefs. In contemporary research, Walton proposed to define a fallacy as “a deceptively bad argument that impedes the progress of a dialogue” (Walton, 1995, p. 256), thereby focusing his definition on both the deceptive and disruptive nature of fallacies. Finally, Tindale (2007) offers a view that is akin to a distinction made by Plantin (1995): fallacies can be deceptive, but they are not necessarily so, since they may arise as a consequence of errors in judgement or appreciation, which are accidental. To resolve the tension, Plantin proposes to distinguish sophisms from paralogisms on the criterion of intention: sophisms are intentional, and thus deliberately deceptive, while paralogisms are accidental, and thus not intended to deceive. All this research points to the fact that fallacies can be deceptive when speaker mean them as such: a corollary of the notion of fallacies by design is that they do not have communicative status—meaning the deceptive intention they try to fulfil must remain covert. Normatively acceptable arguments, conversely, are overt: they are formulated as sincere and acceptable arguments that are meant to convince, and they are precisely meant to be identified as sincere and acceptable arguments that are meant to convince. This is the first notable difference between fallacies and normatively sound arguments.

The second difference, which we relate to the backgrounding/foregrounding dynamics of both types of arguments, has to do with the selectional constraints they operate on information sets. As noted, fallacies by design need to remain covert, which requires them to weaken the chances that a critical context will be mobilised or to make sure, if it is mobilised, that it is an epistemically weak one: we should thus expect them to try to keep critical information as weak as possible (in epistemic terms) and as inaccessible as possible (in terms of processing effort). In other words, deliberate fallacies try to sink the competition. This is not the case with normatively sound arguments: in fact, the best arguments, from a normative perspective, are precisely those that prevail after careful submission to doubt. From a dialectical perspective, furthermore, the best arguments are those that have withstood all potential objections—and this requires systematic and relentless mobilisation of critical information. So, in the end, normatively sound arguments encourage competition and promote critical discussion. This, we contend, translates into diverging selectional constraints on information at the cognitive level.

In this framework, going beyond their normative shortcomings to assess their cognitive operability, we can therefore reinterpret known fallacies as specific strengthening and weakening strategies. Some will focus on strengthening contents in terms of epistemic strength, while others will devote their resources to weakening critical information in terms of its accessibility, and so on. For example, *ad hominem* fallacies reinterpreted in this fashion are weakening strategies meant to decrease the epistemic strength of a piece of information by discrediting its source. Mirroring this dynamic in an opposite direction, *ad verecundiam* fallacies try to implement strengthening strategies so as to increase the epistemic strength of fallacy-friendly information by giving credit to its source. Red herrings would also be weakening strategies, with the difference that what is weakened is the addressee’s likelihood of accessing critical information, as they effectively divert attention away from the argument’s problematic content, and thus away from the representation of relevant critical information.

What could potentially lie ahead of us in this project, then, is a complete reassessment of existing work in fallacy research, which brings together the normative and the rhetorical side of argumentation.

Before we discuss an example of how these constraints might play out in a context of persuasion, we need to add one layer to our rhetorical account of fallacious argumentation, which will take into account the interrelatedness of language, cognition and social context in argumentative contexts (see also Oswald & Herman, 2016). This will eventually lead us to formulate a clearer stance on what rhetorical effectiveness is in the framework we defend.

4.3.3 Pragma-Linguistic Constraints on Argumentation and Rhetorical Effectiveness

The framework outlined above considers that cognitive mechanisms of argumentative evaluation may be disrupted because of the selectional constraints on information typically enforced through the use of fallacies. This, we contend, happens by the operation of strengthening and weakening strategies, which translate into the foregrounding and backgrounding of different sets of information, as just seen. While we so far hinted at the principled ways in which this might happen at the cognitive level, namely through the manipulation of epistemic strength and accessibility, we have not yet gone into the details of how exactly these strategies are realised, notably in terms of their verbal formulation. But looking at them first requires us to think about the different types of inferences that are performed in argumentative evaluation.

At the core of our investigation lies the relationship between two types of inference: pragmatic inference (PI) and argumentative inference (AI). PI is an inference about meaning, which fulfils the interpretative goal of identifying speaker meaning, and which takes as input verbal and contextual information, in order to deliver an output in the form of a representation of speaker meaning. AI is about the acceptability of a justificatory link, which fulfils the evaluative goal of assessing the quality of argumentation, and which takes as input a representation of speaker meaning together with contextual information, in order to deliver an output in the form of an evaluative representation between a conclusion and reasons to accept it (see Oswald, 2018). Both PI and AI converge on one crucial dimension. Even if argumentation theory usually deals with questions related to argument quality and pragmatics is traditionally concerned with issues of meaning, both can indeed be taken to share one concern, namely curiosity for the mechanisms by which representations end up entertained by individuals through communication. Accordingly, in both accounts, quite minimally, the main issue is about explaining why and how a given representation becomes part of the individual's cognitive environment. We have submitted earlier that the cognitive mechanisms invoked to account for PI are similar to those accounting for AI. It is now time to explore this idea more in detail, and we do so by first considering the relationship between PI and AI.

There are good reasons to consider that AI depends on PI. The first of them comes from cognitive considerations. From the above characterisations, it seems indeed that what AI partly takes as input, PI supplies as output—i.e. a representation of speaker meaning. Dan Sperber and his colleagues, in their work on epistemic vigilance, furthermore note that “comprehension of the content communicated by an utterance is a precondition for its acceptance” (Sperber et al., 2010, p. 367). This means that before one can assess whether a claim follows from supporting premises, one needs to be able to represent and understand the content of both the premises and the claim, since very little, if anything at all, is to be gained from an argumentative articulation of meaningless statements.

The second line of inquiry providing evidence in favour of the idea that AI feeds on PI comes from rhetorical scholarship, which has long articulated a claim that rests on the very dependence of AI on PI: the way one frames an argument, in terms of the verbal choices one operates, can have significant impact on its chances of success. Rhetoric, classically construed as the study of the best available means of persuasion (Aristotle, 1995), has been in the business of studying the specificities of formal linguistic features of argumentative messages (syntactic structures, tropes, figures of speech, etc.) since its inception. In his work, Tindale motivates this idea and offers countless examples to show how different formulations might appeal to different audiences, thereby acknowledging the role PI has on AI (see Tindale, 1992, 2007, 2015). In some dialectically-inspired accounts, research on strategic manoeuvring, which incorporates a rhetorical component in the extended pragma-dialectical model of argumentation (van Eemeren, 2010), points to this very idea in its discussion of the notion of *presentational devices*, which are linguistic choices made to enhance the persuasiveness of the message. Very recently, Schumann and her colleagues (Schumann, Zufferey, & Oswald, 2019) offered the first experimental study of the straw man fallacy to specifically cater for some linguistic and pragmatic criteria which influence its acceptability, i.e., its undetectability, showing decisive links between linguistic formulation and argumentative evaluation. All in all, there is overwhelming consensus in argumentation scholarship around the role of language and interpretation on persuasiveness, and thus indirectly on argument evaluation. In principle, hence, indicators converge to establish the relevance of accounts of utterance interpretation for the study of argumentation.

The notion of pragmatic inference and its theorising owes a great deal to Grice's pioneering pragmatic work on meaning (collected in his posthumous 1989 book), as his model of communication as a cooperatively rational conversational undertaking provided the building blocks of contemporary pragmatics. The starting point of his theory is the idea of semantic underdeterminacy, which follows from his (1957) distinction between natural and non-natural meaning. The idea of semantic underdeterminacy rests on the clear-cut distinction between what is said and what is meant. Grice provides a principled account (consisting of a cooperative principle and 4 conversational maxims which further specify what conversational cooperation should amount to) of how conversational participants are able to figure out that what speakers mean is many times different and more specific than what they say. Grice names this type of implicit meaning implicatures. Crucially, in his model, Grice considers that

reaching a (naïve) interpretation of someone's utterance involves inferring speaker meaning, which in turn boils down to recognising (and thereby fulfilling) a communicative intention. Implicatures are thus defined as contents which are implicitly speaker-meant and inferentially hearer-derived. Here we see that the construal of inference offered by Grice is not of an argumentative nature, as PI simply denotes the process by which implicit contents may be reached.

Within relevance theory (Sperber & Wilson, 1995; Wilson & Sperber, 2012), PI is driven by considerations of relevance. That is, what makes addressees appropriately identify speaker meaning is the fact that their cognitive system identifies it as contextually relevant. In other words, understanding someone's utterance amounts to identifying in which way that utterance is relevant in the context at hand. Relevance receives a technical definition, under this cognitive perspective, which characterises it as a relationship between processing effort (the use of cognitive resources to work out meaning) and cognitive effect (roughly speaking, the usefulness of an assumption, which relevance theory defines in epistemic terms). Relevant information, thus, is information that requires little processing efforts and yields significant cognitive effects. This effort-effect ratio is said to be responsible for the selection of information in any cognitive operation: once the system considers that expectations of relevance are fulfilled, it selects the output information and moves on to process other stimuli.⁵

An account such as the one introduced in relevance theory is instrumental to answering our research questions in at least two ways. First, the cognitive nature of the account allows us to characterise fairly precisely what we termed strengthening and weakening strategies: fallacies, in the end, manipulate the perceived relevance of information (relevance being defined as an optimal ratio between epistemic strength and processing effort). Second, using an account of PI to assess AI is an ideal move to assess the strong connection between the persuasiveness of people's argumentations and their linguistic wording, which is of capital rhetorical importance. Earlier in this chapter (Sect. 4.2.2) we set out to explain how the standard treatment of fallacies could be salvaged by offering a rhetorical reinterpretation of its content. We are therefore now able to fully account for this reinterpretation within a rhetoric-pragmatic framework that caters for issues of persuasiveness.

A final note, but an important one, is in order regarding what we mean by *rhetorical effectiveness* when we deal with fallacious arguments. The expression itself is vague, but intentionally so, as it may cover different kinds of disruptive argumentative behaviours. On the one hand, one could simply take rhetorical effectiveness to be a synonym of persuasiveness: rhetorically effective fallacies are those that manage to make sure that their conclusion ends up belonging to the addressee's cognitive environment. This is the Aristotelian sense we mentioned earlier, whereby fallacies give the impression, and only the impression, of having successfully defended the claim. However, we want the expression to cover additional scenarios, such as distractions and failures to spot argumentative infelicities. If a speaker manages, for instance

⁵We refer the reader to additional work in relevance theory (e.g. Sperber & Wilson, 1995; Wilson & Sperber, 2002) for the detailed cognitive account of communicative processes.

through a fallacious *ad hominem* argument, to change the topic of the conversation to her own advantage, even if there is no actual persuasion about the standpoint defended by the *ad hominem*, we still want to call that rhetorically effective. More abstractly, we use the expression *rhetorical effectiveness* to also include scenarios in which a speaker manages to reach a conversational outcome where the fallacious moves she has made remain unidentified. This broader construal of rhetorical effectiveness is fully compatible with our pragmatically-driven account: explaining how information sets are selected and how fallacious arguments constrain this selection indeed affords such a broad construal because it is merely instrumental to explaining a range of different covert argumentative disruptions.

4.4 A Fallacious and Deceptive Erotetic *Tour de Force*

The example we will discuss here comes from the 2009 film *Public Enemies*, directed by Michael Mann and adapted from Bryan Burrough's non-fiction book *Public Enemies: America's Greatest Crime Wave and the Birth of the FBI, 1933–34* (2004).⁶ The book features the stories of many well-known gangsters, among others John Dillinger. In Chap. 9, Burrough recounts how Dillinger has been arrested and is held in Lake County Sheriff's House and Jail, in Crown Point, Lake County, Indiana. Due to rumours about a possible escape, both the prosecutor, Robert Estill, and the female director of that prison, Sheriff Holley, request that Dillinger be transferred to a higher-profile prison, Indiana State Prison, in Michigan City. Dillinger does not want this, as an escape plan is more likely to succeed where he is, in Lake County, than where they want to send him, which is why he instructs his lawyer Louis Piquett to try to talk the judge out of granting the transfer request. We reproduce below the excerpt from the film, whose dialogical format is more practical to analyse.⁷

Context:	Dillinger is shackled to his chair. The courtroom is packed. Walls are lined with deputies holding rifles. Reporters scribble. The gavel quiets the crowd
Piquett (on his feet)	Your honor! Are we to have a hearing in accord with the laws of this nation, or is the State to be permitted to incite an atmosphere of prejudice? The very air reeks with the bloody rancor of malice. The clanging of shackles brings to our minds the dungeons the Czars, not the flag-bedecked of an American courtroom. I request the court to direct that those be removed forthwith!

⁶According to the author, "The conversations and dialogue in this book are taken verbatim from FBI reports, the Karpis transcripts, contemporary news articles, and the memories of participants." (Burrough, 2004, preface). Following this disclaimer, we presume that the dialogue under analysis here has taken place in real life, even if beyond the author's declaration, we have no guaranteed way of ascertaining that it has indeed been reproduced *verbatim*.

⁷Script available at <https://www.imsdb.com/scripts/Public-Enemies.html>, last accessed 27.03.2019.

- Robert Estill: This is a very dangerous man, your Honor.
 Deputy Holley: And I'm responsible for the safeguarding the prisoner.
 Piquett: Who are you?! Are you a lawyer? What right have you to address this court?
- Judge Murray: Alright, remove the handcuffs from the prisoner.
 Robert Estill: Your honor, we'd like to relocate the prisoner. Only Indiana State Prison in Michigan city can guarantee Dillinger will not escape.
- Judge Murray: Sheriff Holley?
 Sheriff Holley: I concur, your honor.
 Piquett: Sheriff Holley, I think it's a very nice jail you have right here. What makes you think there's anything wrong with it?
- Sheriff Holley: There isn't anything wrong with my jail! It's the strongest jail in Indiana.
- Piquett: That's what I thought. But of course, I don't want to embarrass Mrs. Holley. I appreciate that she's a woman and if she's afraid of an escape...
- Sheriff Holley: I'm not afraid of an escape. I can take care of John Dillinger or any other prisoner.
- Judge Murray: Okay, Dillinger will stay here.

Piquett's strategy to get Dillinger to stay in Lake County Jail is rhetorically interesting on a number of counts. First, it instantiates three crucial questions which are meant to take control of the discussion. Second, it instantiates at least two known and possibly three fallacious arguments. Third, it manages to get Piquett's adversaries to do all the argumentative work—so to speak, Sheriff Holley walks into a trap on her own, and she is the one to provide the decisive piece of information that tips the scales in favour of Piquett and Dillinger. Our analysis begins with an assessment of the particular linguistic construction and unfolding of this strategy, in the hope of highlighting how and why the fallacies remain invisible to their addressees until it is too late. For this, we will look at the specific constraints that Piquett's series of questions impose on the interpretation of his words.

4.4.1 A Double Rhetorical Trick Question Realising a False Dilemma

We start with Piquett's opening disjunctive question, which we reproduce in (1) below:

- (1) (a) "Are we to have a hearing in accord with the laws of this nation, or (b) is the State to be permitted to incite an atmosphere of prejudice?"

We take (1) to be an instantiation of the fallacy known as false dilemma (see Boone, 1999; Tomić, 2013), as it indeed seems to us that the terms of the proposed alternative

are neither mutually exclusive nor exhaustive. From a normative perspective, therefore, this move is fallacious. Even if from a formal perspective (1) could be evaluated as deductively valid, the argument fails in that the disjunctive premise underlying the questions is false. We are thus not legitimated in drawing the first disjunct from the disjunction. There is, however, much more to this strategy if we think about how it is likely to be processed.

(1) is uttered at a particular moment and in a particular context: the film stages Dillinger's entrance in the courtroom first through sound (we hear shackles clinking) and then by a vertical pan shot spanning from the feet to the head of the gangster. These are actually the visual and auditory cues that prompt Piquett's first question (1a). Thus, auditory, visual and verbal cues in this case converge to create a context in which Dillinger is represented as being the victim of unfair, borderline unlawful, treatment. Let us incidentally note that the order in which the two disjuncts of the false dilemma are uttered is in itself interesting. While in principle, and all else being equal, a premise such as "either A or B" could be reversed as "B or A" with no notable difference, we submit that this is unlikely to be the case in this context because the order of the clauses seems rhetorically significant. The question contained in (1a) is a rhetorical question whose answer should be ruled out as irrelevant due to its oddness, if not its absurdity.⁸ Yet, we surmise that it is likely to trigger interesting cognitive effects and thus to achieve relevance in a context like the one we just described. The situation itself (a shackled Dillinger, victim of a form of institutional indignity) actually licenses and legitimates the question. While (1b) could be considered as an attack against the institution, (1a) summons what should be standard and normal legal procedure (obeying the laws of the Nation) to confront what Piquett sees as an abnormal situation which does not seem to be compatible with the rule of law. The move orchestrated in (1a) seems to rest on a *modus tollens* that we reconstruct as follows:

- (2) (a) If the court acted in accord of the laws of this Nation, Dillinger would not be shackled.
- (b) Dillinger is shackled.
- (c) Therefore the court did not act in accord of the laws of this Nation.

Since the conclusion is untenable for an institution whose mission is precisely to make sure that the laws of the Nation are observed, it becomes impossible to keep the prisoner shackled or to plead an exception to the application of the law—and the judge indeed complies with Piquett's request.

Regarding the linguistic formulation of this strategy, two important observations are in order. First, if Piquett had opened with (1b), he would have automatically assumed an antagonistic stance, which would have prompted the judge to either defend his court or to sanction the lawyer for the affront. In such a scenario, Piquett would have first supplied his own reading of the situation, and that would have opened his flank to criticism: volunteering one's opinion inevitably comes with the risk of being criticised and refuted. But by expressing his opinion in (1b) instead of

⁸The answer to this question is obviously and unquestionably a very trivial 'yes'.

in (1a), Piquett presents it as a mere description of facts that seems incompatible with his recalling of the court's mission (i.e., the rule of law referred to in (1a)). From a cognitive perspective, this amounts to weakening the likelihood that the judge will engage in criticism and/or sanctioning. Calling attention first to a state of affairs that he deems abnormal under the rule of law makes available a rhetorically interesting option: Piquett is now able to use this as evidence for the relevance of (1b)—and he needs (1b), as we show next.

We venture that our second observation, which concerns the perceived relevance of (1a), makes this line of analysis more plausible. The acceptability of the conclusion (2c) on which (1a) rests is so ethically problematic that Piquett would not have even needed to formulate an alternative. He could have limited himself to requesting the court to remove the shackles based on (1a) only. Yet, (1b) allows him to strengthen the fundamental assumption that the context is unfair to Dillinger. Since the alternative is manifestly absurd (as (1b) is clearly to be rejected), the relevant point is not to know whether the State is allowed to incite an atmosphere of prejudice, but instead to make more accessible, and possibly epistemically stronger, the assumption that seeing Dillinger in shackles is actual evidence of said prejudice. In the immediate aftermath, as he grants the request, it is striking to see that the judge does not invalidate this interpretation: on the contrary, he confirms its relevance, thereby strengthening it. This is why (1b) becomes so difficult to refute.

To summarise: through a double rhetorical trick question, Piquett manages to create a context in which Dillinger is relevantly portrayed as a victim of unfair treatment by a prejudiced State. Additionally, since the latter has manifestly incited prejudice, it becomes difficult for the judge to credibly refute (1b). Piquett has therefore managed to strengthen both the confrontational setting and Dillinger's role of victim (additionally confirmed by Piquett's following utterance "the very air reeks with the bloody rancor of intolerant malice"). This is fundamental in the lawyer's overall rhetorical strategy, because it sets the stage for what comes next: for this we have to keep in mind that whatever follows will be processed against a context in which the court's unfairness and Dillinger's victimhood are highly salient assumptions.

4.4.2 *More Rhetorical Questions to Strengthen the Context*

The deputy's contribution is quickly attacked by Piquett through, again, two rhetorical questions:

- (3) "Are you a lawyer?"
(invited answer: 'obviously not')
- (4) "What right do you have to address this court?"
(invited answer: 'obviously none')

The information-seeking appearance of questions (3) and (4) is manifestly irrelevant, since the answers are obvious. This means that their relevance should be found elsewhere. One option would be to consider that Piquett is out for an *ad hominem*

attack meant to disqualify the speaker as a credible—if authorised—speaker in the courtroom, but we think that the point of this attack actually lies elsewhere.

In terms of their content, these two rhetorical questions are meant to prevent the deputy from addressing the court based on legal considerations (only legally sanctioned individuals may address the court). Thus, Piquett appears, again, through (3) and (4), to make sure that the court observes the laws of the Nation and that Dillinger receives a fair trial. Consequently, (3) and (4) are likely to be perceived as consistent with (1a) and (1b) because they share the same intent. Just as he fought the illegal shackling of Dillinger, he fights to prevent unauthorised individuals from speaking in court. Two different strategies are at play here, one directed at his own *ethos*, the other at the image of the court he wants to convey: while he poses as the defender of the law, the court seems to have contempt for it, in light of its behaviour, which Piquett is very quick to highlight.

In cognitive terms, Pickett is, once again, trying to make sure that whatever happens next is assessed against a context that has strengthened, or foregrounded, the biased and prejudiced nature of the court. This second instantiation of the strategy is cognitively advantageous, simply because it is “more of the same”: yet again, Piquett shows that the court is so hostile towards Dillinger that it allows attendants (in this case, the deputy) to engage in illegal behaviour. And this, as per our cognitive account, is a cognitively optimal move, since it proves easy to process, making that particular set of contextual information highly accessible.

4.4.3 *The Straw Man Question*

The previous sets of questions were preparatory moves for the final touch of Piquett’s erotetic strategy, which is ultimately about manipulating his opponents into admitting that Lake County Jail is a very safe prison. If he manages to do that, then he *ipso facto* refutes the grounds to transfer Dillinger to a more secure prison. Moreover, it would be rhetorically more powerful that this piece of information emerges as an admission from his opponents, rather than as one of his own arguments or refutations. Indeed, if he succeeds, not only will the opponents provide a refutation to their own standpoint; they will also, in the process, look bad for not having worked out their request in a sound way. And for that, he crucially needs to obfuscate the fact that the admission that Lake County Jail is very secure functions as a strong counter-argument to their request for transfer. We will see that he secures that outcome by managing to deflect the function and relevance of that particular admission towards a defence following a character attack on Sheriff Holley. In particular, we will show how Piquett does this through a straw man fallacy, which is a fallacy of refutation in which an arguer’s position is misrepresented in order to be attacked more easily (see e.g., Aikin & Casey, 2011; Lewiński & Oswald, 2013; Oswald & Lewiński, 2014).

In the dialogue, Robert Estill, the prosecutor, states the following:

- (5) “Only Indiana State Prison in Michigan city can guarantee Dillinger will not escape.”

Sheriff Holley, when asked, explicitly concurs, and this is followed by another question by Piquett (6):

- (6) “What makes you think there’s anything wrong with it [Lake County Jail]?”

Through this question, Piquett misrepresents his interlocutors’ position: they have not conceded that there is something wrong with Lake County Jail, but that Indiana State Prison is better, which are two different things. This is a blatant misrepresentation, if only because, topic-wise, Piquett’s question is about Lake County Jail, while the prosecutor’s assertion is about Indiana State Prison. However, we believe—and this is indeed what happens—that Piquett’s straw man has prospects of being rhetorically effective, that is, of disrupting the conversation to achieve the intended result of securing an admission. We can support this analysis through the following five observations.

First, as defended in what precedes, Piquett utters (6) in an immediate context of suspicion of bias against Dillinger. The lawyer has managed to show that the court is prejudiced against the prisoner, which makes the possibility of injustice extremely salient: in such a context, perhaps the prosecution’s request for a prison transfer can even be seen as an unjustified and useless manoeuvre.

Second, (6) is a loaded question which presupposes two important things:

- (7) There is something wrong with Lake County Jail
 (8) The director of Lake County Jail believes that there is something wrong with her prison

These presuppositions are relevant on two counts: first, they echo the hypothesis we just evoked, namely the possibility of injustice and the fact that there might be something fishy behind the transfer request. Nevertheless, we think the rhetorical effect of this presupposition is actually to be found at the conversational level: for obvious face reasons, Sheriff Holley cannot come across as someone who believes that her prison is unsafe, and thus the presupposition in (8) strengthens the need for a defence on her behalf.⁹ She in fact takes on the presuppositional challenge, unfortunately for her without realising that she is actually falling into the trap that will fulfil Piquett’s ulterior motive (to get Dillinger to stay at Lake County Jail). By reacting to (8) through a refutation, she provides Piquett with the admission he was looking for from the beginning.

Third, the strength of the straw man, together with the likelihood that it will not be detected—which, in our view, rests on the interpretation of the straw man as an attack that needs to be countered by Sheriff Holley, and not as a call for an admission—are increased by the attributive format of the question. Piquett uses the second person

⁹The reference to her gender—and, implicitly, to the associated stereotype of inferiority—could function as an additional constraint designed to increase the chances that Sheriff Holley takes on the presupposition and responds to it in a way that makes her admit that her prison is safe. This is all Piquett needs to conclude that there are no grounds whatsoever to transfer Dillinger.

pronoun “you” to address the Sheriff directly. In so doing, he shifts the burden of proof to Sheriff Holley, who finds herself in the position of having to disprove a misattribution for which she is not even responsible.

Fourth, we observe a strange inversion in the dialectical roles due to the misrepresentation Piquett operates in the straw man: at the end of the exchange, the lawyer is in the unnatural position of defending Lake County Jail and in fact appears to be more defensive of this prison than its own director. And this, rhetorically speaking, seems extremely effective, because it casts an unfavourable light on the Sheriff.

Fifth, there is some semantic proximity between what Sheriff Holley says and Piquett’s misrepresentation. The reason for this is that thinking that something is wrong with Lake County Jail could be a possible entailment of considering that only Indiana State Prison can guarantee Dillinger will not escape. The problem, of course, is that this is arguably not a necessary entailment, although the use of the adverb ‘only’ suggests an exclusive reading favouring the entailment. This goes to show that Piquett’s move is not completely off-mark, which is also why it has prospects of being rhetorically effective.

Piquett’s overall argumentative strategy is twofold: on the one hand, his erotetic scheme heavily relies on the audience’s inferential processes to work out the argumentative thread. All questions are either leading, rhetorical or treacherous, and all of them are designed to be rhetorically favourable to Piquett and Dillinger. Some of them target the construction of a context in which he and his client are victims of the court’s unlawful treatment, and others are meant to absolve Piquett from the need to state his position and to formulate complete arguments. On the other hand, these questions are meant to put Piquett’s interlocutors in a defensive position, since in all three cases Piquett attributes a standpoint to his interlocutor, what is more a standpoint that the interlocutors need to exonerate themselves from. The judge has to prevent the crowd from thinking that he is presiding an unlawful court, the deputy needs to disassociate from Piquett’s implication that he is failing to observe the law by addressing the court, and Sheriff Holley needs to fight the misattribution that she believes her prison to be unsafe. Thus, instead of needing to argue, Piquett makes the burden of proof rest on his interlocutors’ shoulders. This strategy culminates with the admission that Lake County Jail is a safe prison, which is all he needed to make his case.

This is a complex erotetic strategy, but we do think that it lends itself to the kind of framework we developed earlier. In cognitive terms, Piquett leads the show because he is able to constrain which pieces of information become relevant in his interlocutor’s minds, as they answer his questions. In (1a) and (1b), he generates expectations of relevance for (i.e., he tries to increase the subsequent accessibility of) any content related to the confrontational and prejudiced attitude of the court and to Dillinger’s assumed victimhood. The context that is then available prominently contains these assumptions. In (3) and (4), he consolidates this context, thereby securing the presence of the same discriminatory assumptions in the cognitive environments of the members of the audience. Finally, in (6), which presupposes (7) and (8), he makes sure that Sheriff Holley interprets the question as an attack, and not as a way of eliciting the admission that Lake County Jail is a very secure prison. He does this by

making the attack interpretation extremely accessible (in light of the first two sets of questions and of the presuppositions). The need for defensive action is thus overwhelming, which is how she ends up unwittingly admitting to what Piquett wanted her to admit.

The rhetorical effectiveness of this erotetic strategy is multifaceted. On the one hand, Piquett has managed to convince everyone that Lake County Jail is safe by getting its director to admit that there is nothing wrong with it. In so doing, he has managed to establish that there is no need to transfer Dillinger. On the other hand, he has successfully managed to conceal the means by which he wanted to achieve that, by distracting the audience into thinking that the dialogue was about a personal attack, when in fact it was about getting Sheriff Holley to volunteer a decisive piece of information. This goes to show that the complexity of the notion of rhetorical effectiveness, as a theoretical construct, stands to gain from the type of account presented here, which captures different subtypes of rhetorical effectiveness, thereby achieving descriptive and explanatory adequacy.

4.5 Conclusion

In this chapter, we have tried to rehabilitate the definition of fallacy given in the standard treatment by showing that what is a problem for the normative question (i.e., its psychological component) is actually an advantage for the explanatory question related to rhetorical effectiveness. We have shown that the cognitive evaluation of arguments can be represented as boiling down to figuring out how compatible sets of information (here the content of the argumentation and the critical context) are, so as to determine whether the conclusion of the argumentation should be included in the cognitive environment of the addressee. We have further established that said compatibility is determined along two cognitive parameters: epistemic strength and accessibility of critical information.

Through our pragma-rhetorical analysis of a complex erotetic argumentative strategy, we have tried to show that our cognitive approach to fallacies has interesting prospects of contributing to argumentation scholarship in novel ways. Our discussion yielded results which we think are indeed methodologically valuable, on at least three different levels. First, from a descriptive vantage point, our account allows us not only to study the specificities of each known fallacy, but also to observe how these specificities might be realised through different linguistic choices. Second, from an explanatory perspective, we are now able to supply a principled—and furthermore testable—account of why these linguistic choices are likely (or unlikely), in a particular context, to be rhetorically successful. Third, we can now start thinking about different types of rhetorical effectiveness, to be documented within one and the same theoretical framework. In terms of the development of the discipline, this pleads for the need for further integrative inquiries developed along the lines introduced here.

References

- Aikin, S. F., & Casey, J. (2011). Straw men, weak men, and hollow men. *Argumentation*, 25(1), 87–105.
- Aristotle. (1955). *On sophistical refutations; On coming-to-be and passing away; On the cosmos* (Vol. 3; E. S. Forster & D. J. Furley, Trans.). Cambridge, MA: Harvard University Press.
- Aristotle. (1995). *The complete works of Aristotle: The revised Oxford translation*. Princeton, NJ: Princeton University Press.
- Bentham, J. (1962). *The handbook of political fallacies* (Vol. 69). New York, NY: Harper & Brothers.
- Boone, D. N. (1999). The cogent reasoning model of informal fallacies. *Informal Logic*, 19(1), 1–19.
- Burrough, B. (2004). *Public enemies: America's greatest crime wave and the birth of the FBI, 1933–34*. London: Penguin Press.
- Correia, V. (2011). Biases and fallacies: The role of motivated irrationality in fallacious reasoning. *Cogency: Journal of Reasoning and Argumentation*, 3(1), 107–126.
- Correia, V. (2014). Biased argumentation and critical thinking. In T. Herman & S. Oswald (Eds.), *Rhétorique et cognition. Perspective théoriques et stratégies persuasives. Rhetoric and Cognition. Theoretical perspectives and persuasive strategies* (pp. 89–110). Bern: Peter Lang.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66(3), 377–388.
- Hamblin, C. (1970). *Fallacies*. London: Methuen.
- Hansen, H. (2015). Fallacies. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2015). Retrieved from <https://plato.stanford.edu/archives/sum2015/entries/fallacies/>.
- Johnson, S. (1996). Fallacies and heuristics. In J. van Benthem, F. van Eemeren, R. Grootendorst, & F. Veltman (Eds.), *Logic and argumentation* (pp. 101–114). Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- Johnson, R. H. (2014). *The rise of informal logic: Essays on argumentation, critical thinking, reasoning and politics* (Vol. 2). Windsor, ON: University of Windsor.
- Johnson, R., & Blair, J. A. (2006). *Logical self-defense*. New York, NY: International Debate Education Association.
- Lewiński, M., & Oswald, S. (2013). When and how do we deal with straw men? A normative and cognitive pragmatic account. *Journal of Pragmatics*, 59, Part B, 164–177. <https://doi.org/10.1016/j.pragma.2013.05.001>.
- Maillat, D., & Oswald, S. (2009). Defining manipulative discourse: The pragmatics of cognitive illusions. *International Review of Pragmatics*, 1(2), 348–370.
- Maillat, D., & Oswald, S. (2011). Constraining context. A pragmatic account of cognitive manipulation. In C. Hart (Ed.), *Critical discourse studies in context and cognition* (Vol. 43, pp. 65–80). Amsterdam: John Benjamins.
- Mercier, H., & Sperber, D. (2009). Intuitive and reflective inferences. In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 149–170). Oxford: Oxford University Press.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34(02), 57–74. <https://doi.org/10.1017/S0140525X10000968>.
- Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Cambridge, MA: Harvard University Press.
- Oswald, S. (2010). *Pragmatics of Uncooperative and Manipulative Communication*. Université de Neuchâtel.
- Oswald, S. (2011). From interpretation to consent: Arguments, beliefs and meaning. *Discourse Studies*, 13(6), 806–814. <https://doi.org/10.1177/1461445611421360e>.
- Oswald, S. (2014). It is easy to miss something you are not looking for: A pragmatic account of covert communicative influence for (critical) discourse analysis. In C. Hart & P. Cap (Eds.), *Contemporary studies in critical discourse analysis* (pp. 97–119). London: Bloomsbury Publishing.

- Oswald, S. (2016). Rhetoric and cognition: Pragmatic constraints on argument processing. In M. Padilla Cruz (Ed.), *Relevance theory: Recent developments, current challenges and future directions* (Vol. 268, pp. 261–285). Amsterdam: John Benjamins.
- Oswald, S. (2018). Pragmatic inference and argumentative inference. In S. Oswald & D. Maillat (Eds.), *Argumentation and Inference: Proceedings of the 2nd European Conference on Argumentation, Fribourg 2017* (Vol. 2, pp. 615–629). London: College Publications.
- Oswald, S., & Herman, T. (2016). Argumentation, conspiracy and the moon: a rhetorical-pragmatic analysis. In M. Danesi & S. Greco (Eds.), *Case studies in discourse analysis* (pp. 295–330). Munich: Lincom Europa.
- Oswald, S., & Lewiński, M. (2014). Pragmatics, cognitive heuristics and the straw man fallacy. In T. Herman & S. Oswald (Eds.), *Rhétorique et cognition: Perspectives théoriques et stratégies persuasives—Rhetoric and cognition: Theoretical perspectives and persuasive strategies* (pp. 313–343). Bern: Peter Lang.
- Plantin, C. (1995). L'argument du paralogisme. *Hermès*, 1, 245–262.
- Schumann, J., Zufferey, S., & Oswald, S. (2019). What makes a straw man acceptable? Three experiments assessing linguistic factors. *Journal of Pragmatics*, 141, 1–15.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., et al. (2010). Epistemic vigilance. *Mind and Language*, 25(4), 359–393. <https://doi.org/10.1111/j.1468-0017.2010.01394.x>.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford: Blackwell.
- Tindale, C. W. (1992). Audiences, relevance, and cognitive environments. *Argumentation*, 6(2), 177–188. <https://doi.org/10.1007/BF00154324>.
- Tindale, C. W. (2007). *Fallacies and argument appraisal*. Cambridge, MA: Cambridge University Press.
- Tindale, C. W. (2015). *The philosophy of argument and audience reception*. New York, NY: Cambridge University Press.
- Tomić, T. (2013). False dilemma: A systematic exposition. *Argumentation*, 27(4), 347–368.
- van Eemeren, F. H. (2001). Fallacies. In F. H. van Eemeren (Ed.), *Crucial concepts in argumentation theory* (pp. 135–164). Amsterdam: Amsterdam University Press.
- van Eemeren, F. H. (2010). *Strategic maneuvering in argumentative discourse: Extending the pragma-dialectical theory of argumentation*. Amsterdam: John Benjamins.
- van Eemeren, F. H. (2015). *Reasonableness and effectiveness in argumentative discourse: Fifty contributions to the development of pragma-dialectics*. Amsterdam: Springer.
- van Eemeren, F. H., Garssen, B., Krabbe, E. C. W., Henkemans, A. F. S., Verheij, B., & Wagemans, J. H. M. (2014). *Handbook of argumentation theory* (2014th ed.). New York, NY: Springer.
- Walton, D. (1995). *A pragmatic theory of fallacy*. Tuscaloosa: University of Alabama Press.
- Walton, D. (2010). Why fallacies appear to be better arguments than they are. *Informal Logic*, 30(2), 159–184.
- Willard, C. (1995). Failures of relevance. In H. V. Hansen & R. C. Pinto (Eds.), *Fallacies. Classical and contemporary readings* (pp. 145–158). Pennsylvania: The Pennsylvania State University Press.
- Wilson, D., & Sperber, D. (2002). Relevance theory. *UCL Working Papers in Linguistics*, 13, 249–287.
- Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge; New York: Cambridge University Press.