

Diagnosing Misattribution of Commitments: A Normative and Pragmatic Model of for Assessing Straw Man



Douglas Walton and Fabrizio Macagno

Abstract This paper builds a nine-step method for determining whether a straw man fallacy has been committed in a given case or not, by starting with some relatively easy textbook cases and moving to more realistic and harder cases. The paper shows how the type of argument associated with the fallacy can be proved to be a fallacy in a normative argumentation model, and then moves on to the practical task of building a hands-on method for applying the model to real examples of argumentation. Insights from linguistic pragmatics are used to distinguish the different pragmatic processes involved in reconstructing what is said and what is meant by an utterance, and to differentiate strong and weak commitments. In particular, the process of interpretation is analyzed in terms of an abductive pattern of reasoning, based on co-textual and contextual information, and assessable through the instruments of argumentation theory.

1 Introduction

This paper has three aims. The first is to model the straw man fallacy as a distinctive form of argument by identifying the argumentation scheme it is based on and by displaying its structure within a normative model of argumentation. This first aim is a task for argumentation theory, and the objective is to show how the scheme is embedded in existing normative models of argumentation accepted in argumentation theory. The second aim is more of an applied task in nature. It is to provide a method for anyone, or more particularly for a student trying to learn how to use methods of

D. Walton (✉)
University of Windsor, Windsor, ON, Canada
e-mail: dalton@uwindsor.ca

F. Macagno
Instituto de Filosofia da Nova (IFILNOVA), Faculdade de Ciências Sociais e Humanas,
Universidade Nova de Lisboa, Lisbon, Portugal
e-mail: fabrizio.macagno@fcsh.unl.pt

informal logic, to apply the theory to real examples of arguments in natural language discourse. The goal here is to be able to take real but fairly simple (so-called easy) cases of arguments where it is suspected that this fallacy may have been committed, and marshal the evidence to determine whether it has been committed or not. The third aim is to provide resources to help a user apply the theory and the method to the so-called hard cases. The examples studied, in this paper and elsewhere in the literature, are mainly political and/or legal arguments. The hard cases can be quite complicated, can be hard fought by both sides in legal or political arenas, and involve interpreting a lengthy corpus of natural language text.

The study begins with a series of easy cases, examples that are simple and short. It is shown that the structure of the argument that is the target of analysis can be represented by a precise argumentation structure in normative argumentation models in these cases. Some of these are textbook cases. While they are “easy” they are also shown to have some hard aspects that can be subtle and tricky to precisely model. However, a method is developed that, it is argued, can deal well enough with these easy cases. It provides an evidential basis for making a sound and well-documented argument that the fallacy has, or has not been committed in the given case.

When the paper comes to the hard cases, the natural language text of discourse in the case exhibits features of a kind that make it more difficult to pin down whether the straw man fallacy has really been committed or not. For example, by wrenching a quote out of its co-text or context, the original speaker’s intentions can be distorted or not represented adequately. In order to address these cases, it is necessary to investigate the pragmatic dimension of a quote, and the pragmatic processes that are involved in the reconstruction of the propositional form of the utterance and its implicit dimension. By distinguishing between such processes and identifying the ones at work in the interpretation of an utterance it is possible to determine the strong commitments of the speaker, and assess whether they have been distorted or not.

The overarching aim of this paper is to rise beyond the narrow approach of seeing the straw man fallacy as simply an embarrassing logical error. In the paper a richer notion of fallacy will be adopted and applied to build a system of fault diagnosis and repair for straw man argumentation. We reconfigure the straw man fallacy so it can be used as a tool to deal with an important general failure in argumentation practices in political argumentation in democratic countries. We present a method that can be used to test an attack on an opponent’s position to see if it meets the normative standards for a rational discussion. If the attacks on an arguer’s position fail the test, the next step is to move to a repair procedure so that the fault can be fixed.

2 Literature Survey

The informal logic textbooks roughly agree on how to define this fallacy, although they do differ on details of the exact wording on how the fallacy should be defined. According to the definition given by (Johnson and Blair 1983, p. 71) the straw man fallacy is committed “when you misrepresent your opponent’s position,

attribute to that person a point of view with a set-up implausibility that you can easily demolish, and then proceed to argue against the set-up version as though it were your opponent's." This definition uses the terms 'position', 'point of view', 'implausibility' and 'demolish'. According to the definition given by Govier (1992, p. 157), the straw man fallacy is committed "when a person misrepresents an argument, theory, or claim, and then, on the basis of that misrepresentation, claims to have refuted the position that he has misinterpreted." This definition uses the terms 'theory', 'claim', 'misrepresentation', 'refuted' and 'position'. According to the definition of Hurley (2003, p. 122) the straw man fallacy is committed "when an arguer distorts an opponent's argument for the purpose of more easily attacking it, demolishes the distorted argument, and then concludes that the opponent's real argument has been demolished". This definition is based on the terms 'distorts', 'attacking', 'demolishes' and 'the opponent's real argument', implying that some argument is attributed to the opponent that is not "real". What should we take these terms to mean? 'Claim' is generally taken to be another word for the conclusion of an argument. The terms, 'attack', 'refute' and 'demolish' seem to be part of a family of terms that are closely related. The terms 'point of view' and 'position' also seem to be closely related.

Bizer et al. (2009) cite an abundant literature showing that the straw man argument is a commonly used technique in political argumentation, and support this claim with examples from current political discourse. Talisse and Aikin (2006, p. 349) claim that the straw man argument is "among the most prevalent forms of fallacious argumentation at work in contemporary political discourse".

Legal cases show how disputes about alleged misrepresentation of speech is a frequent issue disputed in court. If something was misrepresented in a news report, it can constitute an instance of libel. Disputes about how to interpret what was said that attempt to prove or disprove the accuracy of reported speech in a news article can be extensive and hard-fought in such cases. Using an example of an Australian defamation case, Ikeo (2012) shows how the courts try to identify misrepresentation of speech by attempting to reconstruct the interior speech events in order to negotiate the meanings of key expressions that appeared in the reported speech. The factual evidence in such a case is the text of what was actually said, according to documents reporting the wording in the news report.

Walton and Macagno (2010) analyze the fallacy of wrenching from context to show how the technique works by manipulating the meaning of another party's statement through devices such as the use of misquotations, selective quotations and quoting out of context. Using a series of 23 examples, they showed how pervasive this argumentation technique is in political discourse. They analyze wrenching from context as a fallacy that works by unfairly attributing a commitment to another party, a commitment that the other party never really held. Hence it is clear that this fallacy is very closely related to the straw man fallacy. Both work by distorting another party's position or argument, a fault that can only be diagnosed by examining the original text of discourse where that party was quoted, or where some view was attributed to him, and evaluating whether the quotation or statement expressing a view really expresses the arguer's commitment.

According to Walton and Macagno (2010), straw man is often connected to quotation and paraphrasing where the other party's words are changed or distorted to represent a position that is easier to attack. Wrenching from context, in contrast, involves the use of the opponent's exact words, but the context has been altered in order to change the meaning of these words. Also, while straw man always refers to the other party's position in a discussion, wrenching from context can also refer to the position of a third-party, such as that expressed by an expert in an argument from expert opinion. Thus on this view (Walton and Macagno 2010), wrenching from context is a wider strategy consisting in altering the position of one's opponents to support one's own viewpoint. On this view, there are two different argumentation strategies corresponding to these different types of argument and each of them needs to be explained by a different pragma-dialectical role.

Misquotation is a common error in both legal and political argumentation that is closely related to the straw man fallacy (Macagno and Walton 2011). Misquotation is commonly combined with the fallacy of wrenching from context and the straw man fallacy, as shown by an example of cross-examination in the Nuremberg trials (Macagno and Walton 2011, p. 39), and also by a famous example in the trial of Galileo, where his words were twisted around by the prosecutor (Macagno and Walton 2011, p. 39). They also show how misquotation used to distort the other party's position is often closely connected to the *ad hominem* fallacy.

Aikin and Casey (2011) classified three forms of the straw man fallacy they called straw man, weak man and hollow man. In the first form (Aikin and Casey 2011, p. 89), the attacker distorts the other party's position, and uses the distorted version to attribute a significantly less defensible position to her. He then criticizes the distorted position, and draws the conclusion that the other party's position has been refuted. In the second form of attack, the weak man argument, the attacker selects the weakest of several arguments put forward by the other party, attacks these arguments, and then claims to have refuted the other party. In the third form of attack, the hollow man argument, not merely a distorted version, but a complete fabrication of the other party's position is used to refute her.

By analyzing a series of cases of political and legal argumentation, Macagno and Damele (2013) showed how the straw man fallacy can be based (1) not only on distortion of the other party's explicit commitments, but also (2) on distorting the other party's implicit commitments, and even (3) on combining the explicit and implicit commitments to distort the other party's position even more subtly. Using examples, they show why the most powerful and dangerous tactic is the third one. At the same time, it is the subtlest one, and the most difficult to identify and analyze, because it requires the use of Gricean implicature to extract implicit commitments. In *Interpreting straw man argumentation* (2017), Macagno and Walton grounded the assessment of fallacious misattributions of commitments on an argumentation-based model of interpretation. Straw man fallacies were shown to be reports or quotes leading to distorted interpretations of the original utterances. The distortion of an interpretation was represented as an inferred meaning requiring more, and more defeasible inferential steps than the speaker's presumable intended meaning.

Through this model it is possible to reconstruct not only the implicit commitments illegitimately attributed to the speaker, but also explain why this commitment attribution is in fact mischievous.

3 An Easy Example

The following example (Freeman 1998, p. 88) is a very simple and basic instance of the form of argumentation widely taken to represent the straw man fallacy. It has been analyzed in (Walton 2013, p. 255) where it was called the beer and wine example. It is really a pair of arguments. The first argument, attributed to a concerned citizen, states that it would be a good idea to ban advertising beer and wine on radio and television, because these ads encourage teenagers to drink, often with disastrous consequences. The argument seems to be reasonable. It fits the argumentation scheme for argument from negative consequences, a very common form of defeasible argumentation that is in principle reasonable to support a conclusion, provided the premises are acceptable. Both premises would be widely considered to be acceptable to current audiences. It seems likely that such ads would encourage teenagers to drink, barring any evidence to the contrary. And it would be generally accepted that when teenagers drink there are sometimes disastrous consequences. The second argument, put forward by an alcohol industry representative, states that the generalization that you cannot get people to give up drinking, and backs it up with the supportive statement that people have been drinking for thousands of years. This argument, when considered in itself, seems fairly reasonable too. The second argument attacks the first argument, and is designed to refute it. What's wrong with this argument? What's wrong is that it is supposedly an instance of the straw man fallacy.

Despite the differences of terms used to formulate these three definitions of the straw man fallacy, there is a basis of convergence and agreement on how the general *modus operandi* of the fallacy should be described. Johnson and Blair (1983, p. 71) capture the three essential attributes of how the fallacy works as a strategic maneuver of argumentation by formulating the following three requirements. First, the proponent attributes a certain view or position to the respondent. Second, the respondent's real position is not the attributed one, but a different one. Third, the proponent criticizes the attributed position as though it were the one actually held by the respondent. Here we would appear to have three essential requirements for committing the fallacy.

Next, let us turn to the question of how the beer and wine example fits this definition and its essential requirements. Freeman (1998, p. 88) analyzed the argumentation in the example by pointing out that the concerned citizen did not make the claim that it would be a good idea for people to give up drinking. He also pointed out that there is no textual evidence indicating that this claim is the conclusion the concerned citizen was arguing for. By maintaining this in his counter-argument, however, evidence is provided showing that the alcohol industry

representative wants to make it seem to the audience that the concerned citizen is arguing for this claim. Here then all three essential ingredients of the straw man fallacy postulated by Johnson and Blair are present.

What is the straw man fallacy illustrated here, and how can it be proved that it is committed? Some suggestions are given in Freeman's commentary on the example. Freeman (1998, p. 88) distinguished between the following two claims (propositions):

P1: It would be a good idea to ban advertising beer and wine on radio and television.

P2: It would be a good idea to get people to stop drinking.

P1 is the "real" conclusion of the concerned citizen's argument, according to the text of the example that was provided, and the absence of any further text. P2 is the alcohol industry representative's portrayal of that conclusion (according to the text given in the example.)

Freeman asks which is the easier claim to refute, and answers that anyone can see that P2 is much easier to argue against than P1. So here is the explanation of the set-up implausibility cited by Johnson and Blair. This example is a good case to illustrate the basic components of the straw man fallacy and give some beginning shape to how this fallacy needs to be defined and understood. It provides a good point of departure for building a formal model of the fallacy that, as we will show, can be applied to more complex and controversial cases.

However, there is a subtlety to even this simple example that needs to be noticed. The alcohol industry representative's premise stating that people have been drinking for thousands of years would strongly appear to be true, and could be backed up by evidence, such as the ancient Egyptian practice of brewing beer. And it does provide evidence to support his conclusion that you cannot get people to give up drinking, even if this evidence is not conclusive. Moreover, his statement that you cannot get people to give up drinking, the conclusion of his argument, is relevant to the argument of the concerned citizen. For if the generalization that you cannot get people to give up drinking is true, or at least acceptable as a defeasible generalization, it includes teenagers, implying deductively that, since teenagers are people, you cannot get teenagers to give up drinking. So there is a subtle connection between the alcohol industry representative's argument and concerned citizen's argument. The former argument does at least give some basis for attacking the latter argument. It is not absolutely worthless as a refutation.

4 Some Examples That Are Less Easy

In the beer and wine example, the attack on the original argument was claimed to be misdirected because it misidentified the conclusion of the original argument it was attacking (Fig. 1).

But could there also be instances of the straw man argument relating to the two other kinds of argument attacks? Could it be an instance of the straw man

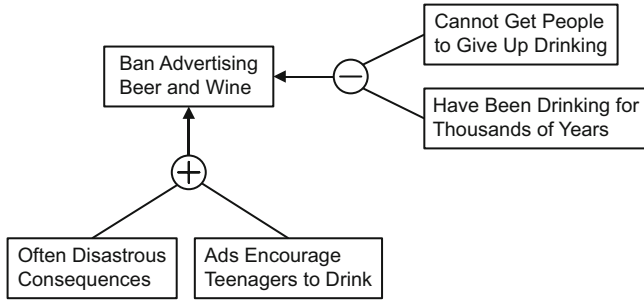


Fig. 1 An attack on a conclusion of a prior argument

fallacy if the attack somehow misrepresented the inferential structure of the original argument, claiming that it fitted an argumentation scheme when really it was a different kind of argument altogether? Or could it be an instance of the straw man fallacy if the attack misrepresented the inferential structure of the original argument because it got one or more premises of the original argument wrong? This latter possibility can be illustrated by another example.

Lewinski and Oswald (2013, p. 168) offered the following argument that they call an “easy” example of the straw man fallacy.

Original Argument: Social policies of the government are plainly inefficient: a number of scientific studies, including one recently published in sociology, expose[d] major faults of the policies.

Attack: It’s funny to say that the government’s social policies are inefficient based on just one scientific study.

In this case the original argument takes the form of the argumentation scheme called argument from expert opinion. However, the attack misrepresented the inferential structure of the original argument because it got one or more premises of the original argument wrong. The premise in the original argument was the statement that number of scientific studies exposed major faults of the policies. The attack criticized the original argument by wrongly assuming that its premise was the statement that one scientific study exposed major faults of the policies.

It is now widely accepted in the artificial intelligence literature that there are three ways to attack an argument (Prakken 2010). You can attack one or more of the premises, you can attack the conclusion, or you can attack the inferential link, the argumentation scheme that joins the premises to the conclusion. An attack on the conclusion of another argument is called a rebuttal. An attack on the inferential link of another argument is called an undercutter. This terminology derives from the well-known distinction made by Pollock between undercutters and rebutters. Pollock (1995, p. 40) drew a distinction between rebutting defeaters, or rebutters, and undercutting defeaters, or undercutters. A rebutter gives a reason for denying a claim. An undercutter casts doubt on whether the claim holds by attacking the inferential link between the claim and the reason supporting it. Pollock used the

following example (Pollock 1995, p. 41) to illustrate his distinction. Suppose that an object looks red to me, but I know that it is illuminated by a red light and red lights can make objects look red when they are not. Knowing this is not a reason for thinking that the object is not red, because red objects look red in red light too. Pollock calls this an undercutting defeater. This example illustrates the defeasibility of arguments from perception.

What has been shown is that straw man arguments can include kinds of arguments based on a misrepresentation of the conclusion of an original argument as well as kinds based on a misrepresentation of a premise of an original argument. Are there also straw man arguments of the third kind, based on a misrepresentation of the argumentation scheme joining the premises to the conclusion? Or at least, if such attempted refutations exist, should they be classified under the heading of the straw man fallacy? This question is open for further research, but another easy example cited by Lewinski and Oswald (2013, p. 168) is relevant.

Original Argument: in fact, the majority voted in favor, but the motion was not accepted since there was no quorum needed for the occasion.

Attack: I'm sad to hear the majority rule does not apply in our Parliament anymore!

In this example, the attacker presupposes that it can be assumed that the generalization or rule on which the original argument was based is an absolutely universal generalization of a kind that is subject to no exceptions. In other words, it is being assumed that the rule is not defeasible. This way of conceiving the rule allows for the following inference to hold: if there was a single case where the majority in Parliament voted in favor of a motion, but the motion was not accepted, then majority rule does not apply in Parliament. This interpretation is of course inappropriate, assuming that everyone understands that there are exceptions to this general rule.

In this case, it would appear that the attack does not fit into either of the first two categories, namely misrepresentation of the conclusion of the original argument or misrepresentation of one or more of its premises. What seems to have gone wrong is that the attacker has misrepresented the nature of the inferential link joining the premises to the conclusion. Instead of representing it as a defeasible inference that is subject to exceptions, he represents it as an absolutely universal generalization of the kind represented by the universal quantifier in logic.

5 Normative Models

Here we outline two current models of rational argumentation, the pragma-dialectical model of van Eemeren and Grootendorst (2004), and the logical argumentation model of Walton (2013). In the logical argumentation model, one of the normative models of dialogue is called a persuasion dialogue. In order for one agent in a persuasion dialogue (called the proponent), to present a successful argument that can rationally persuade the other agent (called the respondent) to

accept its thesis, the argument has to be based exclusively on premises that are commitments of the respondent. In this type of dialogue, if the proponent fails to meet this requirement in an argument it put forward, the argument simply would not count as an instance of a successful persuasion attempt. In this type of dialogue, each party tries to support its own claims using rational arguments. But it is also vitally important that criticisms of the other party's arguments be employed by both sides, including the asking of critical questions and putting forward of counter-arguments that probe into the weaknesses in the arguments of the other side, enabling faults in them to be identified and corrected. In this normative setting, criticizing another argument by merely making it seem that this other argument is based on the commitments of the other agent, when really it is not, is an insidious kind of fault that undermines the whole purpose of the dialogue as a procedure for getting closer to the truth of the matter being discussed.

Thus, for example, whenever you attack an argument put forward by the other side, in order to conform to the requirements of a rational and useful discussion of this kind, you must base your attacking argument on the "real" argument put forward by the other side. When you attack the claim made by the other side as the conclusion of its argument, the argument itself, or attack some part of it such as premises of it, the target of the attack must be a genuine component of the real argument put forward by your co-discussant.

In the logical argumentation model of (Walton 2013), persuasion dialogue is used as a generic category that is meant to include the critical discussion type of dialogue of van Eemeren and Grootendorst (2004). They postulated a list of ten rules required for all reasonable discussions to take part in the type of dialogue they called a critical discussion. Rule 3 (van Eemeren and Grootendorst 2004, p. 191) is called *the standpoint rule*: "attacks on standpoints may not bear on a standpoint that has not actually been put forward by the other party". This rule requires that if the critic attacks an arguer's position, or standpoint as they call it, "it must really relate to the standpoint that is indeed advanced by the protagonist". The failure to meet this requirement, they point out, is contrary to the goal of the critical discussion, because a resolution of the conflict of opinions set in place at the opening stage of a critical discussion is not possible if the attacker distorts the original standpoint.

The general principles in these normative models as formulated above have as an immediate corollary the barring of straw man arguments from a critical discussion, or any other type of persuasion dialogue representing rational argumentation. Lewinski and Oswald (2013, p. 168) stated that their basic pragma-dialectical understanding of the straw man fallacy is that it constitutes a violation of Rule 3 of the critical discussion normative model. Their reason is comparable to the explanation of such a general failure indicated above. They state that contravening Rule 3 "seriously hinders critical testing" because the critic who surreptitiously misrepresents the original argument being attacked hinders the procedure of critical questioning, thereby undermining the ultimate goal of the critical discussion, which is to reach a reasonable resolution of the dispute on the merits of the arguments.

Such arguments violate the general principle formulated above. Let's call it *the principle of rational refutation*, the RR Principle. This general principle can

be formulated as follows: when a proponent in a reasoned dialogue supposed to represent rational argumentation attacks an argument put forward by a respondent, the proponent must always base its attacking argument on the “real” argument actually put forward by the respondent. Let’s call a *refutation* a successful attack that defeats the argument it was directed to. What is being claimed here is that an attack is only a refutation of the argument it was directed to if it conforms to the RR principle.

The RR principle is normative in nature, meaning that it is a high-level principle such that any argumentative discussion can only claim to represent rational argumentation if the attacking arguments used in it conform to the principle. So many might say, fine that’s all well and good, but evidently, judging from the perceived commonality of straw man arguments in such important instances such as political and legal discourse, how can we test arguments to tell whether in real cases the argumentation meets the requirements of the principle or not? Each case needs to be examined on its merits by modeling the individual argument.

The straw man fallacy is very closely tied to the argumentation scheme called argument from commitment. This form of argument comes into play in a kind of situation in which two agents *a* and *b* are engaged in a dialogue, and *a* makes some utterance attributing a commitment to proposition P_1 to *b* in order to support some argument that *a* is putting forward. In the typical case, proposition P_1 is alleged to be a commitment of *b* because in the previous dialogue, *b* went on record as stating P_1 as an assertion. For example, if *a*’s assertion of P_1 was videotaped, or if there are written records of the discussion where *a* clearly asserted P_1 , *a* is said to be committed to P_1 . The evidence of *a*’s commitment to P_1 is the videotape or the written record. This represents the simplest possible type of case, but the problem is that in studying informal fallacies typically we have to deal with the more complex type of case where *a* attributes some other proposition P_2 to *b*, takes it for granted that P_2 either implies P_1 or is equivalent to P_1 , even though P_2 and P_1 maybe worded somewhat differently, and then *a* argues that *b* must be committed to P_2 because *b* went on record as stating P_1 .

The version of the scheme presented in (Walton et al. 2008, pp. 132–135) represents this more complex type of situation. The version of the scheme below is reported accurately, but not quoted, because there are some notational differences.

5.1 Argumentation Scheme 1: Argument from Commitment

Commitment evidence premise	In this case it was shown that <i>a</i> is committed to proposition P_1 , according to the evidence of what <i>a</i> said or did.
Linkage of commitments premise	Generally when an arguer <i>x</i> is committed to P_1 , it can be inferred that <i>x</i> is also committed to P_2 .
Conclusion	In this case <i>a</i> is committed to P_2 .

There are two critical questions attached to this scheme (Walton et al. 2008, p. 335).

CQ1: What evidence in the case supports the claim that *a* is committed to P_1 , and does it include contrary evidence, indicating that *a* might not be committed to P_1 ?

CQ2: Is there room for questioning whether there is an exception, in this case, to the general rule that commitment to P_1 implies commitment to P_2 ?

Any attempt to apply this argumentation scheme to a particular example where it is alleged or suspected that a straw man fallacy has been committed requires searching into the details of the textual evidence in the case. Such a search takes us beyond the abstract normative models that can be applied to easier cases. Now we are taken to the problem of interpreting the wording of a given text to see whether the position attributed to the arguer by the straw man attacker is the same as the arguer's real position. To better appreciate the difficulties in conducting a search of this sort and using it as evidence to support or refute a claim that a straw man fallacy has been committed, let us look at some harder cases.

6 Harder Examples

In 1961, Hugh Trevor-Roper, a British professor of history, wrote a highly negative book review of a book authored by A. J. P. Taylor, another well-known British history professor. The review was so severe that it put Taylor's reputation as a professional historian in jeopardy. Taylor responded by publishing an article entitled 'How to Quote: Exercises for Beginners' (Taylor 1961), which claimed that Trevor-Roper had made up strawman versions of claims supposedly made in his book by distorting and misquoting what Taylor had actually written. Taylor's method of replying to Trevor-Roper's attacks is especially interesting. He drew up a table with two columns and eleven rows (Taylor 1961). Each entry in the left column was an attack on Taylor's book made in Trevor-Roper's review, quoted directly from the review. Matching each of these entries was a corresponding quotation on the right side allegedly marking the passage of the book where Taylor had stated the view that Trevor-Roper attacked. By examining the table, the reader can compare each allegedly distorted version of Taylor's position with his real position as expressed by a word for word quotation of the passage where Taylor had put forward this position in his book.

Not unexpectedly, reading the table made Trevor-Roper's versions of the views he attributed to Taylor look like they did not fairly represent Taylor's real position. In fact the title of Taylor's article, 'How to Quote: Exercises for Beginners', suggested that Taylor had violated the most elementary requirements of serious and respectable historical writing that need to be taught to introductory students taking classes in history. Clearly the article was a serious attack on Trevor-Roper's reputation as a professional historian whose opinions and writings should be taken seriously. The very last line of the article said that Trevor Roper's methods of quotation "might

Table 1 An Entry in Taylor’s Table in ‘How to Quote’

<i>Trevor– Roper (ENCOUNTER, July, 1961)</i>	<i>Taylor (origins of the second world war.)</i>
For what ought the western statesmen to have done when faced by Hitler’s modest demands? According to Mr. Taylor, they should have conceded them all.	Wiser counsels were not lacking. Early in July [1939] count von Schwerin, of the German War Ministry, was in England. He spoke frankly: “Hitler took no account of words, only of deeds.” . . . This advice was disregarded . . . The British statesmen were trying to strike a balance between firmness and conciliation; and, being what they were, inevitably struck the wrong one.

do harm to his reputation as a serious historian, if he had one”(Taylor 1961, p. 73). This argument is a species of personal attack argument well known in the literature as the *tu quoque ad hominem* argument. Trevor-Roper had attacked Taylor’s book in such a serious way that it put Taylor’s reputation as a serious historian in jeopardy, and Taylor responded with this ironic remark implying (1) that it is dubious whether Trevor-Roper ever had a reputation as a serious historian in the first place (2) but that if he did, it has now been ruined by his devious and unfair methods of attack on a colleague violating the basic rules of respectable historical writing.

To give the reader an idea of how Taylor’s method worked, below we have quoted the quotations on both sides in the second row of Taylor’s table (Taylor 1961, p. 72) (Table 1).

In the left box we see an indirect quotation from Trevor-Roper’s book review supposedly representing a view that Taylor expressed in his book, *Origins of the Second World War*. Hitler’s demands are described as “modest”, and it is claimed that Taylor’s position, put forward in his book, is that the Western statesmen should have conceded all of them. The quotation on the right does not describe Hitler’s demands as modest, and it does not state or imply that the Western statesmen should have conceded all of them. Comparing the two quotations certainly makes Trevor-Roper’s argument look bad. It certainly looks like Trevor Roper’s account of what Taylor claimed is a straw man version that has no basis in the textual evidence at all. But that is not the end of the story.

Trevor-Roper wrote in reply (Trevor-Roper 1961, p. 73) that Taylor’s exercises on how to quote were merely calculated to spare him the trouble of argument, so that he did not have to properly answer the points made in Trevor-Roper’s review. He wrote that in his review he was trying to summarize Taylor’s thesis, and argued that a summary of this sort cannot be tied exclusively to single quotations, because it is distilled from many of them. He went on to say that because of the “bewildering inconsistencies” in Taylor’s own presentation of his thesis, it is hard to figure out what he means (Trevor-Roper 1961, p. 73). He claimed that his summaries of Taylor’s view had to be based on combining several passages in the book. By this means he suggested that Taylor’s method of matching one quotation to the other is not a proper way of providing evidence to support the claim that Taylor’s position has been misrepresented. He supported this counterattack by discussing some of the

rows in the table, but not the one quoted above. Trevor-Roper concluded that he had done nothing to be ashamed of.

Rhetorically speaking, however, Trevor-Roper's reply looks to be on the defensive, and comes out as weak. He certainly had a basically valid point that his representations of Taylor's views are summaries. For this reason he was on good ground suggesting that there is something misleading about Taylor's technique of trying to attack the worth of them as interpretations by matching quotation to quotation. The technique looks more decisive than it really is. However, Taylor's method of presenting a table of quotations in this manner does have some points in its favor. It suggests how to deal with disputes about straw man arguments where there are doubts about whether someone's position has been represented in a fair and accurate way by marshaling the evidence on both sides. Surely matching an interpretation of what supposedly represents an arguer's position on an issue with the actual wording quoted directly from the text in which the arguer made the claim at issue is a very good way for moving forward on helping to build a method of addressing suspected instances of the straw man fallacy. So this example is well worth looking at in relation to studying the straw man fallacy.

There is also a more recent hard example that is worth briefly examining. In 2015 a group called SaveTheInternet.in, which lobbies for the equal treatment of all Internet traffic (internet neutrality), used Taylor's method to analyze some text in a discussion paper written by the Telecom Regulatory Authority of India (TRIA).¹ SaveTheInternet.in argued that the authors of a TRIA discussion paper had used "doctoring and cherry-picking" when they presented a text box on page 93 of the discussion paper purporting to be a quotation from the January 31, 2015 issue of *The Economist*.² There were two articles on net neutrality in that issue,³ and SaveTheInternet.in argued that the TRIA authors had "cherry-picked pro-telco statements from both", edited the articles "beyond recognition", "systematically excluded pro-neutrality statements", and had "fraudulently" passed off "their own arguments as The Economist's editorial position".

To back up these criticisms, Save the Internet used the device of a table with two columns, very much in the format of Taylor's table. The column on the left contains quotations from the TRIA discussion paper. The column on the right contained matching quotations from the two articles originally written in *The Economist*. The table used colors to highlight parts of the text. For example, anti-neutrality content that did not appear in the original article but were inserted by TRIA without attribution was highlighted in red. Pro-neutrality content that appeared in the original article but was removed by TRIA without marking the removal was

¹<http://www.trai.gov.in/WriteReaddata/ConsultationPaper/Document/OTT-CP-27032015.pdf>

²https://docs.google.com/file/d/0B_tSiiU5r8nQTwNGemF1Z3FEMmc/edit

³The first article for *The Economist* can be found here: <http://www.economist.com/news/business/21641257-rules-road-internet-will-always-be-work-progress-be-continued>. The second one can be found here: <http://www.economist.com/news/leaders/21641201-why-network-neutrality-such-intractable-problemand-how-solve-it-gordian-net>

highlighted in orange. Anti-net neutrality positions attributed to the telecom industry being fraudulently passed off as the editorial position of were highlighted in pink.

It is difficult to quote one part of the table to show how the technique works in the way that we did with Taylor's table because large chunks of text, according to the Save the Internet table, had been inserted illicitly as representing parts of the text quoted from *The Economist* article. Also, the TRIA discussion paper is 118 pages long, its language is often confusing, and it overwhelms the reader with arguments against neutrality based on outdated information about previous attempts at regulation. So there is a lot of work to do to compare the two texts. Also, the reader really has to look over the original *Economist* articles to be sure that these supposed parts of the text have been illicitly inserted by the TRIA authors. But one telling piece of evidence is that some of the alleged quotations make grammatical errors that would never have made it past *The Economist's* editors. These passages were underlined in the Save the Internet table. To help with the task, the Save the Internet article cites a list of specific points where the text of the TRIA article contains examples of suspicious attributions and arguments.

Lewinski and Oswald (2013, pp. 172–174) also treated a hard example, a case in 2010 where a plane carrying the president of Poland and 95 other high officials and crew members crashed near Smolensk, Russia, killing all on board. There was uncertainty about the cause of the crash in the judicial inquiries, but a conservative Polish party argued that an opposed party were complicit in an assassination plot put together with the Russians. The argumentation on the two sides consisted in contested interpretations of what had been said by the two parties the before the event. One utterance which was particularly contested was ambiguous, and one of the interpretations suggested the possibility of an assassination plot. However the alleged quotations at issue were made in Polish, and therefore there are some linguistic difficulties are involved in translating them in a way that accurately preserves the suspicious implicatures alleged to be involved.

The central problem in this endeavor is how close the correspondence needs to be between the argument attacked by the one agent and the “real” argument actually put forward by the other agent, as identified in the text of discourse in the case. For example, let's assume that the conclusion of the argument attacked is slightly different from the one put forward in the original text, but it is still arguable that it is close enough to it to maintain the force of the attack. This was the problem addressed in (Walton 2013) by proposing building an inference engine that could search through the original text of discourse to put the related propositions together into a package to determine whether there is enough to match to support the criticism of straw man fallacy. It is also possible that the technology to carry out such a task is already available in artificial intelligence systems currently being applied to argumentation problems and tasks. This is a technical problem, but solving it, or even building resources to the point where there is potential for solving it, would be very helpful for providing automated argument assistants that could be helpful for fault diagnosis in argumentation suspected of the straw man fallacy. Working on this technical problem could help with building a method to precisely identify, diagnose and repair arguments in which this fault occurs.

7 A Method for Evaluating Straw Man Arguments

The following sequence is a nine-step procedure for testing an allegation that a straw man fallacy has been committed in a given case. Any such allegation is based on the assumption that we have an example where one argument attacks another. It is also assumed that the argument subject to attack was based on a proposition P_1 claimed to represent the arguer's position. Let's also assume that the arguer's real position P_2 , is taken to be the same as, or equivalent to P_1 .

1. Identify P_1 , the proposition attacked
2. Determine whether P_1 is the conclusion of the argument, one of its premises, or the inferential link (the argumentation scheme) joining the premises to the conclusion.
3. Search for some original text of the argument, such as a quotation, attributable to the arguer that can be used as evidence to compare P_1 and P_2 .
4. If no evidence is available a burden of proof is set in place for the attacker to find a suitable quotation or text that can be used as evidence for P_1 .
5. If the attacker fails to meet this burden of proof, his straw man is defeated.
6. If evidence is available, a comparison needs to be made to determine whether P_1 and P_2 are close enough to equivalence to support the straw man attack.
7. Proving that P_1 is equivalent to P_2 can be carried out by deriving P_1 from the text or quotation containing or indicating P_2 .
8. If the two propositions are not close enough to show equivalence, and the party with the burden can show this by using the textual evidence, a straw man fallacy has been committed.
9. If the two propositions are close enough to show equivalence, as indicated by the evidence, no straw man fallacy has been committed.

Note that the final outcome, whether it is 8 (fallacy committed) or 9 (no fallacy committed), is subject to revision if a different interpretation of the original text can be given that does not support the outcome reached by the procedure. In other words, the procedure represents meta-argumentation that can be closed off, but can re-opened subject to appeal. The reason is that an argument evaluation procedure of this kind as applied to an argument identified in a natural language text of discourse, needs to be regarded as defeasible. The reason for this is that natural language discourse itself is open to pro and con arguments on whether one interpretation is better than another. Nevertheless, the procedure is useful for evaluating a claim that a straw man fallacy has been committed or not any given case, because it is based on evidence.

As indicated in the UML Activity Model of this testing procedure shown in Figure 2, there are three basic ways a straw man fallacy can be committed: (1) The attacker can fail to find and present evidence of P_1 in some quotation or other text attributable to the arguer, (2) The attacker can fail to find and present any evidence showing that some proposition P_2 can be found in the text that is supposedly equivalent to P_1 , or (3) the attacker can fail to use the evidence to show that P_2 is equivalent to P_1 . The notation BoP refers to burden of proof.

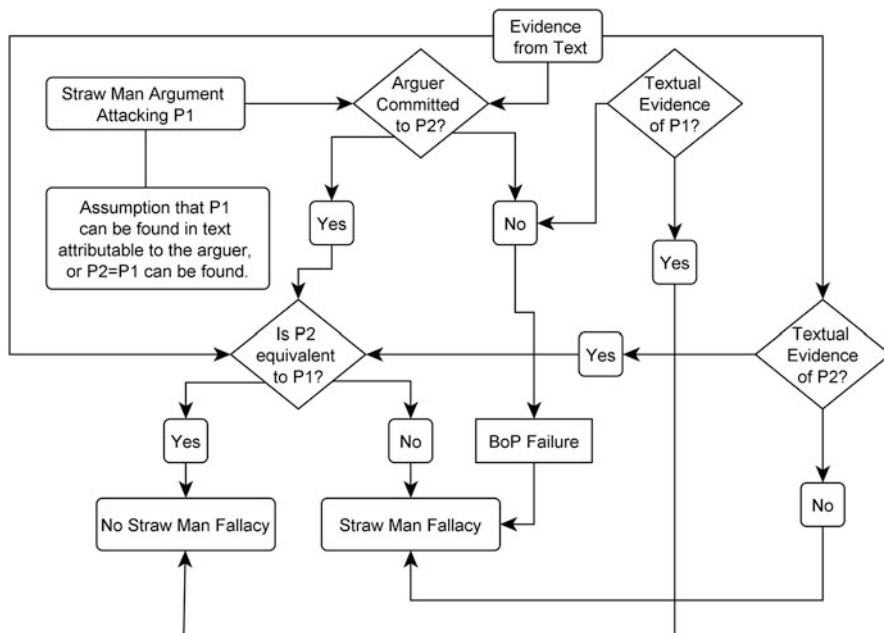


Fig. 2 Activity diagram of the straw man testing procedure

8 A Hard Example That Looks Easy

The problems with applying the test procedure to the hard cases is to are to identify P_2 and prove that P_2 is equivalent (in some appropriate sense) to P_1 . These problems can be illustrated using the Gore and the Internet example. This case was used by Walton and Macagno (2010) to illustrate how an ambiguity in what a speaker actually said can be used to draw out an implicature that makes the speaker appear to have said something different from what he intended by taking the quotation out of context. This case looks like an easy one because it is fairly short. But the explanation needed to show how the argumentation tactic of misrepresentation was employed in this case is quite complex.

The case concerns media reporting of a widely circulated story that then Vice-President Al Gore claimed to have invented the Internet. This claim was taken to be preposterous, and when widely passed around on the media as a joke, it was used an attack to discredit Gore's credibility by making him appear to be a person given to making exaggerated claims. It even suggested that Gore is a liar. Once this accusation had been made and so widely circulated as a joke, an effort was made to try to track down whatever Gore may have said on the subject to see whether the attack was justified or not. Eventually what Gore had said was tracked to an interview with Wolf Blitzer on CNN's *Late Edition* program on March 9, 1999. When Bitzer put the question to Gore to tell what distinguished him from a

challenger for the presidential nomination, Gore gave the following reply (Walton 2013, p. 153).

During my service in the United States Congress, I took the initiative in creating the Internet. I took the initiative in moving forward a whole range of initiatives that have proven to be important to our country's economic growth and environmental protection, improvements in our educational system.

The problem posed then was to try to pin down what Gore could have been reasonably taken to have meant, according to the wording of the above quotation. The question is whether the following conclusion can be inferred from the premise above it.

Premise: I took the initiative in creating the Internet.

Conclusion: I created the Internet.

There are pro and con arguments on whether having stated the premise commits the speaker to having claimed the conclusion. First let's consider a con argument, an argument to the effect that having claimed the premise does not commit the speaker to the conclusion. Gore merely claimed that he took the initiative in creating the Internet, which could mean that he took some kinds of steps that created a political environment that was indirectly supportive to the efforts of the computer scientists at the time that ultimately led to the creation of the Internet.

To support this interpretation, one could argue that taking the initiative to do something does not imply actually doing it yourself. Consider the statement that President Franklin D. Roosevelt took the initiative in creating the atomic bomb. We all know that he did not create the atomic bomb by himself, or even take part in the building of it alongside the scientists and engineers. But we also know that it required his support to back the continuing work on building the nuclear weapon. So it could reasonably enough be claimed that he took the initiative in creating the bomb without drawing the conclusion that he created the bomb. But there are also pro arguments (Walton 2013, p. 154). Consider the statement that in 1902 President Theodore Roosevelt took the initiative in opening the International Court of Arbitration at The Hague. The making of this statement strongly suggests that it is reasonable to infer commitment to the claim that Roosevelt opened the International Court of Arbitration at The Hague.

So this is a hard example in which it is difficult to support or refute the allegation that Gore's opponents used a straw man against him. It depends on answering the following question. Was the statement that Gore claimed to have invented the Internet based on a reasonable interpretation of what he actually said in the interview or not? The best that can be done is to offer arguments pro and con, based on differing interpretations of the text. This is OK, but the problem posed is to see if the method can be extended further to deal better with such cases. The problem with such hard cases is the inference from the interpretation of P_2 , what was said, is not based on deductive necessity, but on pragmatic principles at work in interpreting an utterance. On this perspective, the heart of the problems in hard cases of straw man lies in the means of establishing what was said, and the means of reconstructing what was meant by the utterance.

9 What Is Meant, What Is Said, and Speaker's Commitments

The possibility and the problem of distorting or misrepresenting another's view are inherently dependent on the divergence between the linguistic content and what a speaker means by using it (Carston 2002, p. 15; Grice 1975). The clearest and most evident cases are when "what is said is not even a part of what is meant, but is merely a vehicle for conveying what is meant" (such as in cases of irony), or cases in which "what is said is included in what is meant, but constitutes only a small part of what is meant" (such as in indirect answers) (Carston 2002, p. 16).

However, the semantic representation of an utterance always needs to be pragmatically enriched in order to reconstruct the proposition expressed (Sperber and Wilson 1995, pp. 179–180; Wilson and Sperber 2004). For example, referents need to be assigned to indexicals (Bezuidenhout 1997); the conceptual content of linguistic expressions need to be narrowed or broadened, or their scope specified; linguistic constituents need to be added in order to form a complete proposition; ambiguous expressions need to be disambiguated (Capone 2009; Carston 2002, p. 27). When the proposition expressed is determined, further implicatures can be drawn (Grice 1975), namely forms of inference characterized by cancellability (an utterance made in different contexts can lead to different implicatures) calculability (considering the speaker's compliance with pragmatic principles) and independent functioning (implicatures have propositional forms distinct from the explicature and function independently of the explicature as the premises and conclusions of arguments) (Carston 1988, pp. 156–158). For this reason, the disparity between what the linguistic expression was used to say and what is meant by it corresponds to two discrepancies at different levels (Carston 1988, p. 155, 2002, p. 15):

1. Between the meaning of the linguistic expression used and the proposition explicitly expressed by the utterance thereof (i.e. what is said);
2. Between what is said and the intended implications (the implicatures or implicit import) of an utterance.

In this sense, the encoded linguistic meaning underdetermines what is said, and what is said underdetermines what is meant (Carston 2002, p. 19). The passage from the linguistic meaning to what is meant is determined pragmatically. The explicature, namely the proposition expressed, and the implicatures of an utterance are pragmatic phenomena, is based not only on the linguistic meaning of an utterance but also on the context, the co-text, and the background assumptions of the interlocutors (Soames 2002, p. 79).

The problem of determining what the speaker meant by making his utterance is crucially linked with the problem of establishing his commitments and the strength there of. This issue can be addressed considering the accessibility of the propositions that can be attributed to the speaker as part of his meaning. The force of commitments can be analyzed by taking into account both the nature and the content of the information constituting speaker's meaning.

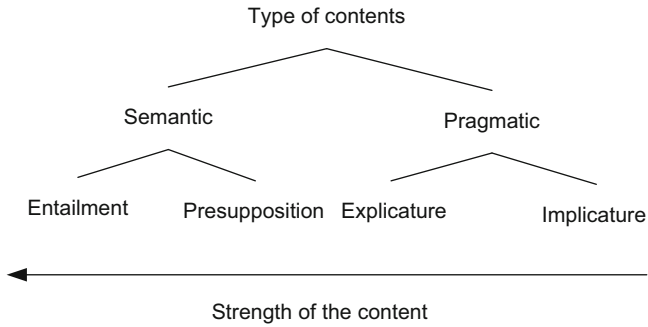


Fig. 3 Strength of contents and strength of commitments

From a theoretical point of view, the strength of commitments depends on the type of content derived from an utterance. Entailments, explicatures, presuppositions, and implicatures constitute the meaning of an utterance. However, entailments and presuppositions are the byproduct of semantic inferences, while explicatures and implicatures result from pragmatic ones. As Moeschler put it, semantic inferences are stronger than the pragmatic ones; for this reason, semantic entailments and presuppositions yield stronger propositions than pragmatic explicatures and implicatures, in turn having different strengths (Sperber and Wilson 1986). Since “the stronger the inferred content is, the more confident is the audience about the speaker’s commitment,” it is possible to conclude that “whereas a speaker cannot deny a semantic inference without contradiction, in the case of a pragmatic inference, she can correct her explicature, and also deny her implicature without contradiction”(Moeschler 2013). The degree of strength of content and the corresponding strength of commitment can be represented as shown in Fig. 3 (Moeschler 2013):

Clearly this model is purely theoretical and based on the nature of the content. In practice these levels are blurred, as the propositional form is determined also through pragmatics (explicatures), and depending on the evidence provided by the context, implicatures can become hard to cancel or retract (Capone 2009, p. 60). For example we consider the following assertion, uttered by an academic (A) to a colleague of the same area (B):

1. I am weak in statistics too.

By using the adverb “too”, the speaker can refer to the interlocutor or other unspecified colleagues in general. In absence of a more specific context of conversation, B can report this utterance claiming that A intended to accuse B of being weak in statistics. Would this indirect report count as straw man? By failing to project his intentions strongly, the speaker opens up the possibility of being attributed a strong commitment. However, considering that such an attribution depends on an explicature that is not supported by clear contextual evidence, B is entering a

dangerous area. In order to analyze the boundaries of the attribution of commitments resulting from pragmatic processes, it is necessary to take into account the problems in determining the implicit contents.

The determination of the commitments resulting from pragmatic processes is essentially bound to the notions of accessibility and relevance. On this view, an implicit content needs to be accessible, namely immediately retrievable within a given context or relevant background knowledge, and relevant. Both requirements are basically dependent on the degree that a given content contributes to the conversation. We can address this issue by considering the problem of determining the possible commitments resulting from the implicatures and implications of an utterance. Typical conversations are characterized by implied contents that are not fully determinate, namely (Soames 2002, p. 84):

cases in which there are (descriptively enriched) propositions q_1, \dots, q_n (which are not necessary consequences of p), such that it is determinate that the speaker's utterance is an assertion of each q_i , and there are other (descriptively enriched) propositions q^*_1, \dots, q^*_n (also not necessary consequences of p), such that for each q_i^* it is indeterminate whether or not the speaker's utterance is an assertion of it.

A clear example is the following assertion, uttered by a professor to a graduate student in the philosophy department (Soames 2002, p. 83):

2. Carl Hempel lived on Lake Lane in Princeton.

The professor intended to convey the meaning that 2* the famous philosopher of science Carl Hempel lived on Lake Lane in Princeton. However, other inferences can be drawn from this assertion, namely that 2**, a former member of the Vienna Circle lived on Lake Lane in Princeton, etc. All these inferences can be accessible given the broader context (an academic conversation). However, while the first inference can be reasonably expected to be drawn by the hearer, the other ones can be drawn, but they do not contribute to the conversation, namely it is not relevant thereto. It is a weaker inference that does not result in the speaker's commitment thereto (Moeschler 2012). In this sense, the existence and the force of a commitment is directly bound to the notion of relevance (Soames 2002, p. 79):

In order for p to be asserted by an utterance of a sentence, it is not enough that conversational participants be in possession of information which, together with the speaker's utterance, might, after long or careful consideration, support an inference to p . Rather, the speaker must have reason to believe both that p is a potentially direct, immediate, and relevant inference for all conversational participants, and that the conversational participants recognize this belief of the speaker.

In order for a proposition to become part of the speaker's commitments, it needs to be "something the relevance of which to the conversation is potentially obvious to all" (Soames 2002, p. 79). For example, we analyze the following utterances, made by a woman to her husband, conveying distinct types of commitments:

- A. Buy me a precious, woman necklace, and give it to me as a present.
- B. Buy me a necklace.
- C. Why don't you buy me a necklace?

- D. It is a nice necklace. I would really love one.
- E. There is a nice necklace in the shop at 5, Melbourne Street.
- F. This necklace is made out of gold.

In A, the speaker is directly and strongly committed to the content that he wants the interlocutor to buy her a precious, woman necklace (semantic content). In B, the commitment is similar; however, the speaker is not directly committed to the fact that the necklace needs to be precious and for women (explicatures). In C, the request is indirect (indirect speech act, conventional), and even more in D and E, where the requests are drawn from unconventional implicatures (maxims of manner). In particular, the strength of the implicature in E is weaker than the one in D, given the same context, as it is less accessible than the latter (it implicates D by the maxim of relevance). Finally, in F the speaker is committed to the fact that the object she is referring to is gold, but she can be only weakly or not committed at all to other possible propositions, such as that the necklace is expensive, or precious, or she would like one, etc. In this latter case, the propositions that can be derived in the specific context are indeterminate and for this reason result in no commitments.

These distinctions between strengths of contents and levels of commitments are useful for establishing what a speaker can be held to endorse and be accountable for. In order for a commitment to be attributable to the speaker, it is necessary to establish whether it results from a strong intention (Capone 2009), which is made explicit (and retrievable by the interpreter or quoter) through textual and contextual evidence (Capone 2012, 2013; Ifantidou 2001; Macagno 2017; Macagno and Capone 2016). In this sense, even though quotations mostly represent only an illocutionary act and the proposition expressed, selecting, marking, and emphasizing the dimension thereof more convenient to the quoter's interests (Clark and Gerrig 1990, p. 779), they need to take into account the conversational dimension and the context. In this sense, a test for the reasonableness of a quotation would be the fulfillment of a burden of quotation (Macagno 2016; Macagno and Capone 2016). The quoter, responsible for the selection of the aspects reported (Clark and Gerrig 1990, p. 792), should be able to prove that the direct or indirect quote (which is often the result of a pragmatic processing of the utterance) is supported by the context. In this sense, the representation (or "demonstration") of the speaker's commitments can be conceived as a conclusion of an implicit interpretive argument, grounded on textual, contextual, and conversational evidence. The speaker is held responsible for the contents directly resulting from his strong communicative intention, namely the ones relevant to the conversation, supported through a pattern of reasoning the implicit conclusion that constitutes the goal of such a discourse or part of discourse (Lascarides et al. 1996; van Eemeren and Grootendorst 2004; Walton 1989, 2003, 2004). In case of weak or non-strong commitments, resulting from pragmatic processes that are not overtly connected with his strong communicative intention (namely not clearly relevant in the specific context), the quotation can be extremely dangerous and fallacious, unless adequately backed by reasons fulfilling a burden of quotation.

The reconstruction, assessment, and justification of a quotation or representation of a speaker's commitments can be represented using patterns of argumentative reasoning (Walton et al. 2008). The fact that an utterance can be interpreted in different ways, leading to distinct communicative effects, does not mean that the two constructions are equally reasonable. The commitments expressed by an utterance (corresponding to the illocutionary force, its propositional meaning, and its implicit content) can be analyzed in terms of presumptions (Levinson 2000) that need to be assessed together with other presumptions also of different kind, resulting from the co-text and the communicative context (Macagno and Walton 2013). On this view, interpretation becomes the conclusion of an argumentative process of reasoning grounded on the abductive pattern of reasoning from best explanation (Walton et al. 2008):

9.1 *Argumentation Scheme 2: Reasoning from Best Explanation*

Premise 1	F (an utterance) is an observed event.
Premise 2	E_1 (Interpretation 1) is a satisfactory interpretation of the meaning of F .
Premise 3	No alternative interpretations given so far ($E_2, E_3 \dots E_n$) are as satisfactory as E_1 .
Conclusion	Therefore, E_1 is a plausible hypothesis, based on what is known so far.

This pattern of reasoning can be assessed using a set of critical questions, namely criteria for evaluating dialogically the reasonableness and acceptability of an inference. In this specific case, the critical questions are the following ones (Walton et al. 2008):

- CQ1: How satisfactory is E itself as an interpretation of F , apart from the alternative interpretations available so far in the dialogue?
- CQ2: How much better an interpretation is E than the alternative interpretations ($E_2, E_3 \dots E_n$) given so far in the dialogue?
- CQ3: How far has the dialogue progressed?
- CQ4: Would it be better to continue the analysis of the interpretation of F further, instead of drawing a conclusion at this point?

The possibility of straw manning a speaker or simply misrepresenting his commitments resulting from his utterances is based on the coexistence of two possibly satisfactory interpretations of his utterance(s) F . An interpretation, however, is not sufficient to explain F , but to explain F based on reasons, and more specifically on reasons stronger than the ones supporting alternative interpretations (premise 3). In this fashion, even harder cases, such as the ones consisting in reporting metaphorical utterances (Macagno 2016), can be explained and reconstructed using patterns of argumentative reasoning.

The criteria of strength of commitment and reasonableness (or burden) of quotation can be used to analyze the attributions and misattributions of intentions to the speaker. For example we analyze the hard case mentioned above:

I took the initiative in creating the Internet.

This assertion was made in a specific co-text: Al Gore was talking about his service in the United States Congress, and pointed out that he “took the initiative in moving forward a whole range of initiatives.” The quote “I created the Internet” or the indirect report “Gore said that he created the Internet” can be analyzed by considering the nature and the reasonableness thereof. First, the quote is not an exact quote; it results from a subtractive explicature, eliminating “I took the initiative of.” This explicature, inasmuch as a pragmatic process, needs to be grounded on textual evidence. Since Al Gore is talking about his service in the US Congress and his initiatives, it is reasonable that the elements subtracted contribute to the issue of describing the importance of his *political initiatives*. Moreover, the whole purpose of his speech was to point out that he *moved forward* important initiatives, of which the Internet was the most important one. The proposition attributed to Gore, “I created the Internet” is the result of a pragmatic processing of his utterance based on an intention that is not obvious to the addressees. Such a reconstruction, in other words, cannot be supported by the evidence found in the speech. At best, the utterance could be analyzed as ambiguous, but in this case the quoted had to fulfill a burden of selecting and depicting an ambiguous utterance. He had to have and provide evident reasons for reconstructing a communicative intention that is not obvious (Macagno 2016), resulting in weaker commitments.

10 Conclusions

Three goals were identified at the beginning of this paper: (1) to model the straw man fallacy as a distinctive form of argument based on an argumentation scheme is based on its structure in a normative model of argumentation, (2) to provide a method for an argument analyst to apply the theory to real examples of an argument in natural language discourse, such as when used in politics or law, to identify cases where this fallacy has been committed, and (3) to provide resources to help a user apply the theory and the method to hard cases. In these hard cases, interpreting the natural language text of discourse where the fallacy has supposedly been committed exhibits complexities of a kind that make it more difficult to pin down whether the straw man fallacy has really been committed or not.

The first two goals were achieved by using the easy cases to show that the testing procedure method provides a method for making a sound and evidence-based argument that can be applied to determining whether the straw man fallacy has been committed in the given case. The third goal was to provide resources to help a user apply the theory and the method to the so-called hard cases. In these cases, the natural language of the text of discourse where the fallacy has supposedly been

committed exhibited complexities of a kind that made it more difficult to pin down whether the straw man fallacy has really been committed or not. To this purpose, the pragmatic processes involved in reconstructing what is said and what is meant by an utterance have been described and analyzed. Strong commitments, resulting from the propositional form of an utterance, its entailments, and its presuppositions, correspond to a stronger representation of the speaker's intentions, and thus result in stronger commitments. Explicatures and implicatures are used as the sources of weaker commitments. In these latter cases, a deeper analysis of the context and co-text is needed to ascertain the real communicative intentions and provide the best explanation of the speaker's utterance. The distinction between strong and weak commitments allows one to establish an implicit "burden of quotation" onto the quoter.

All three goals were achieved, but it was shown that the problems posed by trying to achieve the third depend on resources that have not yet been not fully enough developed to be fully adequate for this purpose. Further research is needed. In particular, the distinction between stronger and weaker commitments remains at a purely theoretical level of development. In practice – as shown in the paper – the implicit pragmatic dimension intrudes into the semantic one. Moreover, the idea of "burden of quotation" needs to be further and better defined by providing precise rules and conditions. Even so, because the paper does provide resources to help move forward with the third task, it succeeds in its aim of not only modeling the straw man fallacy in the setting of the current normative models of argumentation, but also moves forward to fulfilling its secondary aim. This is the aim of building an outline of a well-developed method to enable a user to apply the theory to realistic examples of argumentation of significant interest, such as arguments used in politics, law, and other natural language settings. Further case studies of hard examples are needed to additionally test the theory and refine it.

Acknowledgements The research for this paper was funded by Insight Grant 435-2012-0104 from the Social Science and Humanities Research Council of Canada and by the Investigador FCT Grant no. IF/00945/2013 from the Fundação para a Ciência e Tecnologia of Portugal.

References

- Aikin, S. F., & Casey, J. (2011). Straw Men, Weak Men, and Hollow Men. *Argumentation*, 25(1), 87–105. <https://doi.org/10.1007/s10503-010-9199-y>
- Bezuidenhout, A. (1997). Pragmatics, semantic undetermination and the referential/attributional distinction. *Mind*, 106(423), 375–409. <https://doi.org/10.1093/mind/106.423.375>
- Bizer, G., Kozak, S., & Holterman, L. A. (2009). The persuasiveness of the straw man rhetorical technique. *Social Influence*, 4(3), 216–230. <https://doi.org/10.1080/15534510802598152>
- Capone, A. (2009). Are explicatures cancellable? Toward a theory of the speaker's intentionality. *Intercultural Pragmatics*, 6(1), 55–83. <https://doi.org/10.1515/IPRG.2009.003>
- Capone, A. (2012). Indirect reports as language games. *Pragmatics & Cognition*, 20(3), 593–613. <https://doi.org/10.1075/pc.20.3.07cap>

- Capone, A. (2013). The pragmatics of indirect reports and slurring. In A. Capone, F. Lo Piparo, & M. Carapezza (Eds.), *Perspectives on Linguistic Pragmatics* (pp. 153–183). Heidelberg: Springer. https://doi.org/10.1007/978-3-319-01014-4_6
- Carston, R. (1988). Implicature, explicature, and truth-theoretic semantics. In R. Kempson (Ed.), *Mental Representations: The Interface between Language and Reality*, (pp. 155–181). Cambridge: Cambridge University Press.
- Carston, R. (2002). *Thoughts and utterances: the pragmatics of explicit communication*. Oxford: Blackwell Publishing Ltd.
- Clark, H., & Gerrig, R. (1990). Quotations as Demonstrations. *Language*, 66(4), 764. <https://doi.org/10.2307/414729>
- Freeman, J. (1998). *Thinking Logically* (Prentice H). Englewood Cliffs.
- Govier, T. (1992). *A Practical Study of Argument*. Belmont: Wadsworth.
- Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Academic Press.
- Hurley, P. (2003). *A concise introduction to logic*. Belmont: Wadsworth.
- Ifantidou, E. (2001). *Evidentials and Relevance*. Amsterdam: John Benjamins Publishing Company.
- Ikeo, R. (2012). Misleading speech report in the media with a special reference to an Australian defamation case. *Journal of Pragmatics*, 44(10), 1183–1205. <https://doi.org/10.1016/j.pragma.2012.05.003>
- Johnson, R., & Blair, A. (1983). *Logical self-defence*. Toronto: McGraw-Hill Ryerson.
- Lascarides, A., Copestake, A., & Briscoe, T. (1996). Ambiguity and Coherence. *Journal of Semantics*, 13, 41–65. <https://doi.org/10.1093/jos/13.1.41>
- Levinson, S. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA.: MIT Press.
- Lewiński, M., & Oswald, S. (2013). When and how do we deal with straw men? A normative and cognitive pragmatic account. *Journal of Pragmatics*, 59, 164–177. <https://doi.org/10.1016/j.pragma.2013.05.001>
- Macagno, F. (2016). Reporting and Interpreting Intentions in Defamation Law. In A. Capone, F. Kiefer, & F. Lo Piparo (Eds.), *Indirect Reports and Pragmatics* (pp. 593–619). Cham: Springer. <https://doi.org/10.1007/978-3-319-21395-8>
- Macagno, F. (2017). Defaults and inferences in interpretation. *Journal of Pragmatics*, 117, 280–290. <https://doi.org/10.1016/j.pragma.2017.06.005>
- Macagno, F., & Capone, A. (2016). Interpretative disputes, explicatures, and argumentative reasoning. *Argumentation*, 30(4), 399–422. <https://doi.org/10.1007/s10503-015-9347-5>
- Macagno, F., & Damele, G. (2013). The Dialogical Force of Implicit Premises. Presumptions in Enthymemes. *Informal Logic*, 33(3), 361. <https://doi.org/10.22329/il.v33i3.3679>
- Macagno, F., & Walton, D. (2011). Quotations and Presumptions-Dialogical Effects of Misquotations. *Informal Logic*, 31(1), 27–55.
- Macagno, F., & Walton, D. (2013). Implicatures as forms of argument. In A. Capone, F. Lo Piparo, & M. Carapezza (Eds.), *Perspectives on pragmatics and philosophy* (pp. 203–225). Cham: Springer International Publishing.
- Macagno, F., & Walton, D. (2017). *Interpreting Straw Man Argumentation. The Pragmatics of Quotation and Reporting*. Amsterdam: Springer.
- Moeschler, J. (2012). Conversational and conventional implicatures. In H.-J. Schmid (Ed.), *Cognitive Pragmatics* (pp. 407–434). Berlin: Mouton de Gruyter.
- Moeschler, J. (2013). Is a speaker-based pragmatics possible? Or how can a hearer infer a speaker's commitment? *Journal of Pragmatics*, 48(1), 84–97. <https://doi.org/10.1016/j.pragma.2012.11.019>
- Pollock, J. (1995). *Cognitive Carpentry*. Cambridge, Mass.: MIT Press.
- Prakken, H. (2010). An abstract framework for argumentation with structured arguments. *Argument & Computation*, 1(2), 93–124. <https://doi.org/10.1080/19462160903564592>
- Soames, S. (2002). *Beyond rigidity: The unfinished semantic agenda of naming and necessity*. Oxford: Oxford University Press.

- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition*. Oxford: Blackwell Publishing Ltd.
- Talisse, R., & Aikin, S. F. (2006). Two forms of the Straw Man. *Argumentation*, 20(3), 345–352. <https://doi.org/10.1007/s10503-006-9017-8>
- Taylor, A. J. P. (1961). How to Quote: Exercises for Beginners. *Encounter*, 71–73.
- Trevor-Roper, H. (1961). A Reply. *Encounter*, 73–74.
- van Eemeren, F., & Grootendorst, R. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge: Cambridge University Press.
- Walton, D. (1989). *Informal logic*. New York: Cambridge University Press.
- Walton, D. (2003). Defining conditional relevance using linked arguments and argumentation schemes: a commentary on professor Callen's article, rationality and relevancy: conditional relevancy and constrained resources. *Michigan State Law Review*, 4(4), 1305–1314.
- Walton, D. (2004). *Relevance in argumentation*. Mahwah, NJ: Lawrence Erlbaum Publishers.
- Walton, D. (2013). *Methods of Argumentation*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139600187>
- Walton, D., & Macagno, F. (2010). Wrenching from context: The manipulation of commitments. *Argumentation*, 24(3), 283–317. <https://doi.org/10.1007/s10503-009-9157-8>
- Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation Schemes*. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511802034>
- Wilson, D., & Sperber, D. (2004). Relevance theory. In L. Horn & G. Ward (Eds.), *Handbook of Pragmatics* (pp. 607–632). Oxford: Blackwell. <https://doi.org/10.1016/j.pragma.2009.09.021>