Tweeting fallacies: An exploratory study on fallacy accusations on Twitter

José Ángel Gascón

Universidad de Murcia

Abstract: The fallacy approach to argument pedagogy has been criticized as being overtly critical, theoretically defective and encouraging an adversarial attitude. In order to solve some of those issues, the effects of fallacy teaching on the arguer's behavior should be studied empirically. Here I present an exploratory study in which I take a look at how accusations of fallacies are made on Twitter. 865 accusations were analyzed according to seven criteria: (1) whether the fallacy is identified, (2) whether it is *mis*identified, (3) whether the accusation was substantiated, (4) whether the substantiation makes reference to the context, (5) whether the accuser relies on the "taxonomic technique", (6) whether the accuser relies on a problematic theory, and (7) whether the accuser is willing to discuss the accusation. Both the findings of the study and the reliability of the criteria are discussed.

Keywords: common fallacies, empirical research, fallacies, fallacy accusations, informal logic, on-line argumentation, Twitter.

1. Introduction

Since the beginnings of informal logic, the concept of fallacy has taken a central place in the study and pedagogy of argumentation. Taxonomies and characterizations of fallacies were the main pedagogical approach in early handbooks such as Ralph Johnson and Anthony Blair's *Logical Self-Defense* or Kahane's *Logic and Contemporary Rhetoric*. One of the most important purposes (I am even tempted to say *the most important purpose*) of some well-known theories of argument, such as Douglas Walton's logical approach and pragma-dialectics, is to provide a proper account of fallacies. The claim that knowledge of the fallacies is a cornerstone of argumentation and critical thinking instruction is nowadays almost a triviality.

From time to time there have been, however, certain dissenting voices. Gerald Massey (1975, 1981) argued that there can be no theory of bad arguments, since there is an asymmetry between good and bad argumentation: whereas we can prove that an argument is valid by showing how it instantiates a valid pattern in a certain logical system, in order to prove its invalidity we would need to show that it is invalid in any logical system whatsoever--even those not yet known. The problem with Massey's objection is that it relied on formal logic as a theory of argument appraisal, but several informal logicians have also questioned the concept of fallacy and its pedagogical usefulness.

Maurice Finocchiaro (1981) is an important case in point. According to him, "the concept of a fallacy as a type of common but logically incorrect argument is a chimera" (17), since the arguments usually presented as examples of fallacies in textbooks are either not common or correct. Instead, textbooks usually misinterpret arguments, exaggerating their intended strength, to make them fit in the categories of fallacies. Along

similar lines, Maarten Boudry, Fabio Paglieri, and Massimo Pigliucci (2015) argued that characterizations of fallacies are not particularly useful to detect flawed arguments due to what they call the "fallacy fork": on the one hand, if those characterizations pick out common patterns of argument, then it turns out that many instances of such patterns are cogent arguments; on the other hand, if they manage to identify only flawed arguments, then they are not commonly found in real-life argumentation. In other words: either fallacies so characterized are not intrinsically flawed arguments or they are caricatures of what we find in real discussions.

Regarding the effects of teaching fallacies, Blair has held that "it can be useful, in certain circumstances, when done appropriately" (1995: 328). He admits that, if one relies on poor and oversimplified conceptions of fallacies, teaching them will almost certainly do "more harm than good" (335). Nevertheless, when properly taught, fallacies can be pedagogically useful and "are obviously parts of learning how to evaluate arguments"[1] (333). However, David Hitchcock (1995) disagrees: "everything that can be said with the use of these [fallacy] labels can be said without them, and in general said more clearly" (325). He argues that teaching fallacies fosters an attitude of looking for "a basis on which to convict and sentence" arguments, rather than making a genuine effort to understand them (326). And, finally, Catherine Hundleby (2010) has shown that current textbooks on argumentation and critical thinking promote precisely such an adversarial attitude. Students are taught outdated and poor conceptions of fallacies and they are not encouraged to seriously discuss the merits and deficiencies of arguments, but rather to dismiss them simply because they fit in a fallacious pattern.

Independently of one's position in this discussion, one aspect of it that we should notice is that many of the disputed claims are *empirical*. It is an empirical question whether fallacies are commonly committed, whether knowledge of the fallacies improves argumentative skills, whether it encourages superficial and oversimplified criticism or whether it fosters an adversarial attitude. Unfortunately, however, to my knowledge very few empirical studies have been conducted specifically to test those claims. Some works purporting to provide empirical evidence, such as Gary Jason (1987) on whether fallacies are common or Finocchiaro (1987) on what kinds of fallacies there are, merely rely on particular debates or texts--two presidential debates, in the case of Jason, and Galileo's *Dialogue Concerning the Two Chief World Systems*, in the case of Finocchiaro. There have also been quite a few studies on the effects of critical thinking courses--with rather depressing findings (Willingham 2008)--but none focusing specifically on fallacies. The studies conducted by Frans H. van Eemeren, Bart Garssen, and Bert Meuffels (2009) are the closest we have to empirical research on fallacies, but they focus on whether the pragma-dialectical rules have conventional validity--i.e., whether ordinary arguers endorse them--rather than on how arguers use the categories of fallacies.

I am afraid that gap will not be filled in this article, since I lack the necessary knowledge and skills (and patience) to conduct a series of empirical studies that could shed some light on the effects of teaching fallacies. Instead, here I present what may be called an initial exploration, a glance at how arguers who know about fallacies criticize arguments. I analyzed accusations of having committed a fallacy on the social network Twitter in order to see how they are made. Obviously, we do not know whether the accusers have attended a course about fallacies or have learnt about them elsewhere, and

---

[1] Recently, however, Blair (2023) argued that fallacies should *not* be taught to undergraduates.

if so, where and how. Therefore, we will be hardly able to draw any conclusion about the effects of teaching fallacies or to assess the depth of people's understanding of fallacies. Nevertheless, this study can give us an initial hint on how the concept of fallacy is used by ordinary arguers. This may be of interest, I believe, because even though the concept of fallacy (in its argumentative sense) is becoming increasingly popular, it is not yet part of our everyday vocabulary (Hitchcock 1995: 325). Hence, if that concept is used by an arguer, that indicates that she has been in contact with *some* pedagogical text on fallacies--even if it is just Wikipedia. My goal in this study was to collect some information about how people who have had that contact accuse others of having committed a fallacy.

2. Methodology

Twitter[2] was chosen as the medium from which to collect accusations of fallacies. The advantages of this social network for my purposes is that its users frequently engage in discussions (forming "threads" of messages) that are recorded and publicly accessible. 2080 "tweets" (i.e., messages) containing the word "fallacy" were retrieved on 7 July 2022 using the on-line tool Vicinitas.[3] The sample of 2080 was then cleaned up and analyzed by myself. This must be acknowledged as a weakness of the study, in which ideally several analysts should have been used. For this reason, in the following pages I will be as specific as possible about how the criteria were applied. Future studies should remedy this problem by using more analysts and checking for inter-analyst reliability.

Given the lack of specificity of the search term and the nature of Twitter, many of the results were irrelevant. Hence, as a first step, the sample was cleaned up by removing those tweets that should not be considered for some of the following reasons:

- Repeated tweets ("retweets"): 558.
- Automated messages (bots): 63.
- Tweets in foreign languages: 19.
- Jokes: 13.
- Other reasons: 173.

Repeated tweets were clearly marked by the letters "RT" in the spreadsheet that Vicinitas generated. Of all the repetitions of each tweet, only one of them was included in the sample that was analyzed, and the rest of the repetitions were discarded. Messages by bots were easily spotted because (apart from the fact that they were nonsensical) they were not part of a conversation with other users and they ended with a series of three symbols. The last category, "other reasons," refers to reasons for exclusion that were too specific and not frequent enough to merit a category on their own. The category includes tweets that were part of a thread that was analyzed as a whole (and whose first tweet was also in the spreadsheet and subsequently analyzed), cases in which the word "fallacy" appeared in a quotation from someone else, cases in which the name of the user contained the word "fallacy," cases in which the user simply recommended a show or a blog post whose title contained the word "fallacy," tweets that were telling a fictional story or poetry, and cases in which the tweet was just too incoherent and could not be understood.

---

[2] Now renamed as "X".

[3] https://www.vicinitas.io

826 results were thus excluded, leaving a sample of 1254 items of genuine instances of use of the concept of fallacy. However, here I am focusing on accusations of *argumentative* fallacies, that is, of having produced an argument that is flawed in some respect, whereas that sample contained different uses of the term "fallacy." Some cases referred to accusations of being *biased* rather than of having presented a fallacy. It was necessary to exclude these due to the ambiguity of concepts such as "sunk cost fallacy," which more often than not referred to a psychological phenomenon and not an explicit argument. Others were accusations of having made a *false claim*, rather than of having put forward a fallacious argument. This is the less academic sense of "fallacy," which is often also considered more widespread than the argumentative sense--although, as can be seen in the figures below, this is far from obvious. Others were theoretical reflections, in which nobody is accused of having committed a fallacy, but the tweet is a comment with theoretical interest only, as in: "I believe the *ad hominem* fallacy is the most frequent fallacy." And, finally, others were replies to accusations of having committed a fallacy.

- Falsities: 199.
- Replies to accusations: 85.
- Bias: 57.
- Theoretical reflections: 48.

Hence, in this second step, 389 results were excluded, which leaves us with 865 genuine accusations of argumentative fallacy to be analyzed. Even though the proper analysis of the data has not yet begun, this result may already give us some insight into how people use the term "fallacy."[4] Fig. 1 shows the corresponding percentages of each use of the term.
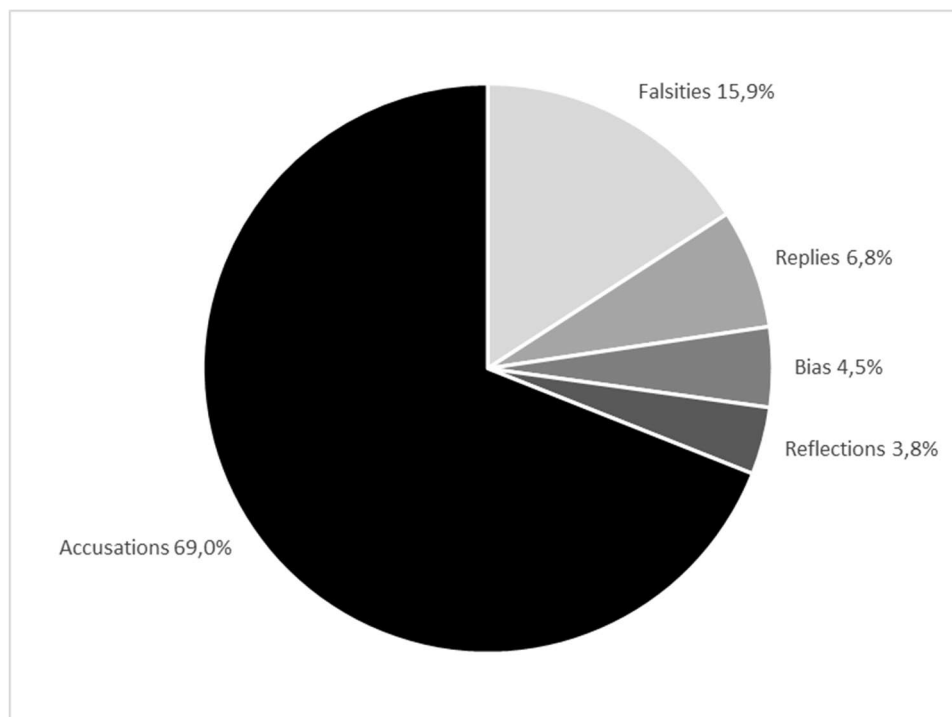

Figure 1. Uses of the term "fallacy"

---

[4] I thank an anonymous reviewer for suggesting this.

Interestingly, as can be seen here, the use of the term "fallacy" to refer to flawed arguments seems to be the most frequent (69%). Even though argumentation theorists sometimes assume that the term is most commonly used with the meaning of "false claim," this data shows that such a sense of "fallacy" is much less frequent (15,9%). Ordinary people seem to have adopted the argumentative sense of the term.

Each of the 865 tweets was analyzed according to seven criteria, which were formulated as yes-no questions. A third value, "Dubious", was also used in those cases in which it was too difficult--even taking the words and the surrounding context of the tweet into account--to provide an answer for the criterion. The criteria were chosen because they point to what can arguably be considered as indications of a reasonable accusation of fallacy. For certain criteria, the ideal answer would be "Yes," whereas for other criteria it would be "No". The criteria are the following.

Criterion 1: Is the kind of fallacy identified? When someone accuses an arguer of having committed a fallacy, at the very least the accuser is expected to specify *which* fallacy has been committed. As Finocchiaro puts it (1987: 266): "when one characterizes an argument as a fallacy (in the literal sense), he must be able to specify what special kind of fallacy it is, because otherwise the claim only means that there is something wrong with the argument."

That claim, however, although it seemed initially plausible when this study began, can be put into doubt. Johnson and Blair warn us that (1994: 56): "the objective in learning the fallacy approach is *not* to enable you to throw these labels around in casual conversation to impress people with your logical savvy." In fact, we will see in the next section that there are reasons to believe that accusations of *unidentified* fallacies might turn out to be more informative, since the arguer often explains the fallacy instead of naming it. But that will have to wait until we discuss the results.

An accusation was assigned the value "Yes" in this criterion only when a specific label was provided, such as "*ad hominem*," "*ad ignorantiam*," "strawman," or whatever. Sometimes, however, accusations did not include an explicit label but rather an explanation of the fallacy--such as "you are attacking the arguer instead of the argument." In those cases, even though the explanation usually made clear what fallacy they were talking about, the value "No" was assigned to criterion 1. Instead, the explanation was considered as a substantiation (see criterion 3 below). The reason for this is the purpose of differentiating between using category labels and using explanations in accusations of fallacies.

Criterion 2: Is the fallacy misidentified? This criterion refers both to accusations in which the accuser correctly spots a fallacy but uses the wrong label--such as accusing someone of committing a strawman when the arguer actually committed an *ad hominem*--and to those in which the arguer did not commit any fallacy at all. It also includes labels of fallacies that are not recognized in the studies of argumentation and do not fit into current conceptions of fallacies--for instance, "false assertion fallacy."

This criterion is tricky, as it heavily relies on the analyst's own evaluation of the alleged fallacy. For this reason, special care was taken to be charitable *to the accuser*. Only when it was absolutely clear that the identification was wrong was this criterion answered in the affirmative. When there was even the slightest possibility that the accuser could be (even partly) right, the answer to this criterion was "No". Hence, a "No" in this criterion does not mean that I agree that there is a fallacy there; it means only that

considering the argument as a fallacy is not an absurd or manifestly wrong interpretation. On the other hand, sometimes the thread was missing (deleted tweets or suspended Twitter accounts) and I could not check the allegedly fallacious tweet, so in those cases I assigned the value "Dubious" to this criterion.

Criterion 3: Is the accusation substantiated? This is one of the most significant criteria used in this study, since it is easy to see whether an accusation fulfils it or not, and also because it undoubtedly marks an important characteristic of reasonable accusations of fallacies: they must be explained or justified. Many argumentation scholars have insisted on this point. An accusation of fallacy obviously involves an evaluation of an argument, and, as Finocchiaro (2007: 254) points out, "argument evaluation can be done seriously only if one gives reasons supporting the evaluative claim." More specifically, Christopher Tindale (2007: 13) writes that "we need to learn not just how to identify fallacies but also to explain clearly what is fallacious about them." And Walton (1995: 33) holds that an accusation of fallacy "needs to be backed up by certain kinds of evidence that meet a burden of proof appropriate for such an allegation." For an affirmative answer to this criterion, *any* kind of substantiation was considered enough--explanation of the fallacious pattern, of the context, an analogy or whatever.

Tweets are famously limited to 280 characters, which might make it difficult for users to properly justify their accusations of fallacies. For this reason, the analysis was not restricted to isolated tweets, but rather for each tweet the whole conversation in the thread was checked. If a user made an accusation in a tweet and justified it in a later tweet within the same thread, the accusation was considered justified. The same methodology was applied to all the other criteria--for instance, if a user made an accusation and only later identified the fallacy, criterion 1 was considered fulfilled.

Criterion 4: Does the substantiation make reference to the context? The possibility of answering this criterion with a "Yes" obviously depends on an affirmative answer to the previous one. If the accusation was not substantiated in the first place, the answer to this criterion will be "No".

In informal logic, reference to context is generally considered necessary in order to assess an argument. Identification of fallacies without making any reference to the context is one of the problems that Hundleby (2010: 286) detects in the examples offered by textbooks: "examples manufactured by the textbook author eliminate further contextual elements that can lend credibility to a competing argument." A good accusation of fallacy, then, is supposed to take into account the broader context. *Any* reference to the context (the circumstances of the arguer or the audience, contextual information related to the content of the argument, or whatever) was considered enough for an affirmative answer to this criterion.

Criterion 5: Does the accusation rely on the taxonomic technique? The taxonomic technique is the assumption that "the presence of an argument scheme that *may* be fallacious always makes an argument fallacious" (Hundleby 2010: 287). That is, it is a method of identifying fallacies that relies solely on whether a particular argument instantiates a fallacious scheme. Such a view is currently regarded as oversimplifying and wrong by most argumentation scholars. Stephen Toulmin, Richard Rieke, and Allan Janik put it as follows (1984: 131):

> …arguments that are fallacious in one context may turn out to be sound in another context. Therefore, we shall not be able to identify any intrinsically fallacious forms

of arguing. Instead we shall try to indicate why certain kinds of argument are, in practice, fallacious in one or another kind of context.

An accusation of fallacy was considered as relying on the taxonomic technique when the accuser explained the fallacious scheme instead of the precise flaws of the particular argument. Sometimes the accuser simply included a link to a webpage where the allegedly committed fallacy was explained. And, finally, depending on the context of the accusation, the value for this criterion could also be "Yes" when the accuser just named the fallacy without any explanation or justification.

In the absence of evidence that the accuser relied on the taxonomic technique, the value for this criterion was "No" by default. Furthermore, for certain kinds of fallacies--such as strawman, cherry-picking, red herring, or *non-sequitur*--there is no specific fallacious scheme, so in those cases the accuser cannot rely on the taxonomic technique. And, of course, if the fallacy is not identified (criterion 1), then the value for this criterion will also be "No".

Criterion 6: Does the accusation rely on a problematic theory? This idea of a "problematic theory" may be problematic in itself, but the criterion was introduced in order to identify accusers' reliance on conceptions of fallacies that are manifestly at odds with (traditional or modern) scholarship. An example is the use of wrong names of fallacies, such as "fallacy of *reductio ad absurdum.*"[5] Another example is the presence of very questionable claims in the substantiation of the accusation, such as "common sense is a fallacy" or "all appeals to authority are fallacious" (this latter case also involves taxonomic technique, which is a kind of problematic theory).

The difficulty in analyzing this criterion is that many accusations do not make manifest the theory on which they rely--especially when it is a bare, unsubstantiated accusation. In the absence of any clear sign of a problematic theory, the value for this criterion was "No." As we will see, there were also several kinds of fallacies that do not fit easily into current lists of fallacies in the argumentation scholarship but that cannot be said to be at odds with them; in those cases, I assigned them the value "Dubious."

Criterion 7: Is the accuser open to discuss the accusation? Finally, an accusation of fallacy can be either a definitive judgment or an initial criticism of the argument that could be discussed. Johnson and Blair emphasize the latter use (1994: 57):

> …the charge of fallacy is nothing more than an initial critical probe of the argument. It is an attempt to locate a possible weakness, not the bold (and sometimes arrogant) assertion that because of this flaw, the argument is worthless. Even if the charge of fallacy is justified in a given instance, that does not mean the argument cannot be repaired over the flaw. Nor does it follow that the conclusion of the argument is false.

Therefore, the issue here is not whether the accuser is willing to continue the discussion with the accused, but whether the accuser is willing to discuss the fallacy accusation specifically. Sometimes this can be seen clearly--when, for instance, the accuser "blocks" the other user on Twitter or says that he or she will not discuss the matter further. But other clues were considered as well. An accuser was also considered unwilling to discuss the accusation when she or he made a bare accusation without substantiating it or referred to hidden interests or bias on the part of the arguer. On the other hand, an accuser was

---

[5] Apparently, many people on the social media believe that *reductio* is a kind of fallacy which resembles the strawman fallacy.

considered open to discussion when the accusation was well substantiated, when the accuser gave reasons or asked questions together with the accusation, or when the accusation was formulated without too much confidence ("I think", "there might be").

Admittedly, criteria 2, 5, and 6 are the most difficult to apply and their results might be less reliable. Criterion 2 (misidentification) depends on an examination of the thread (which was not always possible) and on the analyst's judgment. The application of criterion 5 (taxonomic technique) was based on clues of the accuser's beliefs about the concept of fallacy, which can be very tricky, and the value "No" was charitably assigned when no clues were found. And a similar interpretive problem affects criterion 6 (problematic theory), together with the fact that the analyst's conception of what is "problematic" might be challenged by other argumentation scholars. Criteria 1, 3, 4, and 7, on the other hand, seem more reliable, also because there was no default value in the absence of clues, so that a "No" in them will be as significant as a "Yes".

To conclude the study, after the analysis of the 865 tweets according to these criteria, the labels of fallacies used in the accusations (when they were identified) were collected in a spreadsheet. That will give us an overview of the kinds of fallacies that were most frequently mentioned.

3. Results and discussion

3.1. Characteristics of the accusations

Let us begin then by showing the results of the application of the criteria. I will give them in percentages rather than absolute values (see Fig. 2).
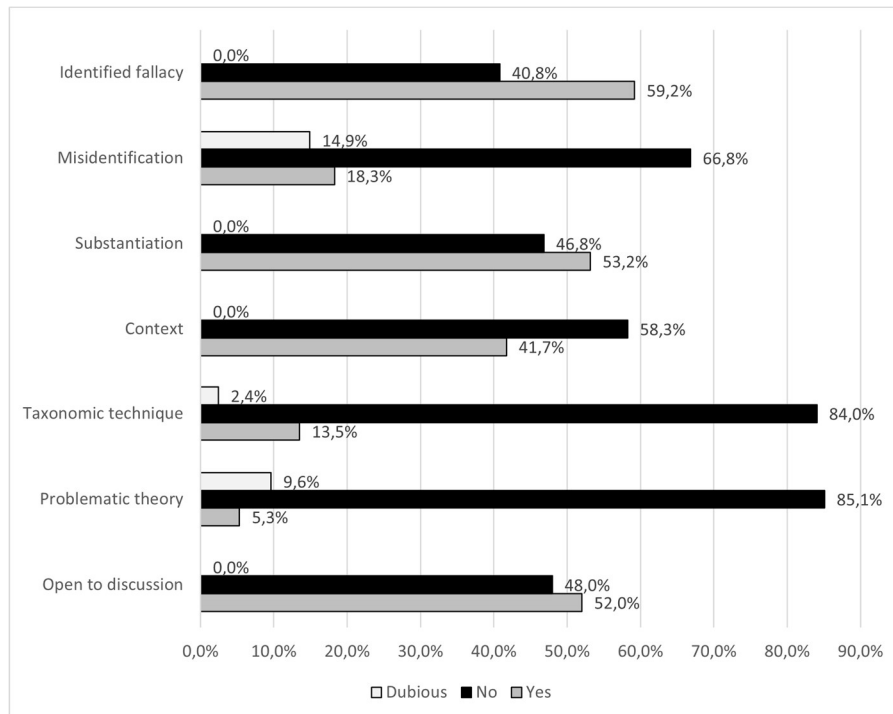

Figure 2. Analysis of accusations

Only three criteria have dubious instances: misidentification (14,9%), taxonomic technique (2,4%) and problematic theory (9,6%). These were the criteria that were based

on interpretive clues and evaluation of the accused tweet and, for this reason, sometimes it was not possible to determine the value with confidence. Misidentification stands out with dubious instances because, in order to apply that criterion, it was necessary to compare the accusation with the argument that was being accused of being fallacious, and in many cases those tweets had disappeared. When that was the case, I assigned the value "Dubious" to it. In the case of problematic theory, the value "Dubious" is also high because, as I explained in the previous section, it was assigned to fallacy labels that did not fit with any list of fallacies in the scholarship but that did not contradict current scholarship either.

More than half of the accusations (59,2%) identified the fallacy, while the rest (40,8%) simply pointed out that there was a fallacy, without adding a label to it. For the purposes of deciding whether the fallacy was identified and whether the identification was wrong, and also of collecting the kinds of fallacies mentioned, some laxity was exercised with the labels of fallacies. So, for example, the following tweets were considered as correctly identifying the fallacy that I have put in square brackets:[6]

Nice apples to oranges comparison fallacy there. [False analogy]

That's just you lying again. No opinions were offered. Gableman wasn't discredited. Name calling is a fallacy. [*Ad hominem*]

Ah, good to see the 'No true Conservative/Communist' fallacy in the wild. [No true Scotsman]

That's another example of the bifurcation fallacy. [False dilemma]

Certain tweets included accusations of having committed more than one fallacy. In some extreme cases, the accusation included a whole set of fallacies. For example:

Illogical fallacies you make with your implied statements:

1. Black or white fallacy.

2. Genetic fallacy.

3. Composition/division fallacy.

4. Burden of proof fallacy.

5. Appeal to emotion fallacy.

6. Slippery slope fallacy.

7. The Texas sharpshooter fallacy.

8. Anecdotal fallacy.

And also:

Look up these fallacies:

Gross generalization.

Guilt by association.

Categorical Fallacy.

---

[6] Some of the tweets quoted in this paper have been slightly edited for orthographical and grammatical correction without altering the contents.

> Insufficient Sample.
>
> Begging the Question.
>
> *Ad Hominem*.
>
> You've just committed all of them.

In those cases, *all* the fallacy labels were included in our collection of mentioned fallacies, which will be discussed in the next subsection. Obviously, when so many fallacies are mentioned in a single accusation, the accuser is unlikely to explain or substantiate them--even considering the whole conversation in the thread.

Among the accusations that did not specify the fallacy that was allegedly committed, we find the following:

> That's not evidence, that's just more fallacy.
>
> I'm sure there's a logical fallacy somewhere but I'm just tired. Why?
>
> That's a common fallacy for people who think they're clever and don't understand the English language.
>
> I think they cover this sort of logical fallacy in philosophy class before they show you where to hang your coat and stuff.
>
> No Beth! As people are saying here the 14M people's mandate belongs to the Conservative Party, NOT Boris Johnson personally. Britain is not a Presidential political system and Boris Johnson is not President, he is a Prime Minister. It annoys me this fallacy is being promoted.

The first four examples might give the impression that accusations in which the fallacy is not identified tend to be very poor--without substantiation and closed to discussion. However, as the fifth example shows, an unidentified accusation can be quite detailed and explanatory. I will come back to this in my discussion of criterion 3 (substantiation).

Given that an accusation was regarded as a misidentification only when the error was manifest, it is no surprise that the percentage of misidentifications was relatively low (18,3%). The following are some examples of accusations considered as misidentifying a fallacy:

> Indeed. Ad absurdum is an argumentative fallacy you rely on a great deal, unfortunately, in your bad faith baseline characterizations of those you disagree with.
>
> "All B is A" is not the corollary of "All A is B." That's the fallacy of composition.
>
> That is an equivocation fallacy. You're taking two unrelated events and clicking them together as relatable. The feelings from Cloud are always unspecified. You're taking a book from 2021 and matching it with one from 20 years ago.
>
> Name fallacy I'm afraid.
>
> Fallacy of the excluded middle. I don't know his real views or attitude but he plays as if he is both on tv in exchange for ratings and lots of money. Don't begrudge a man a living. Lots of you would read his script for $10 million+ a year as well.

Reductio ad absurdum is not a fallacy but a well-known logical pattern. The formal fallacy of inferring "All A is B" from "All B is A" can be called "illicit transposition" or perhaps "illicit conversion"--as has been called by other users--but certainly not "composition" in any of the senses of this term that I am aware of. "Equivocation" refers to the use of an expression with different meanings, not to a false analogy. "Name fallacy"

is not a fallacy that I have been able to locate either in argumentation textbooks or in webpages on argumentation and fallacies. And, of course, "excluded middle" refers to a law of classical logical systems, not to a fallacy consisting on disregarding a third possibility in a dilemma.

Regarding criterion 3, a little more than half of the tweets (53,2%) justified their accusation. Many of those who did not were bare accusations such as the following:

> If your only argument is a logical fallacy, you have no argument.

> Strawman fallacy.

> Motte & Bailey fallacy.

> Your fallacy is cherry picking.

> Logic fail. Fallacy of the single cause.

> Appeal to emotion logical fallacy detected and argument disregarded.

> Nice fallacy.

> The claim begs the question, asserts non-sequiturs, and is the 'no true Scotsman' fallacy.

There were, however, also many accusations that were supported by remarkably good justifications. The following are some examples (I have included comments in square brackets when some context was necessary):

> The argument that Boris has a mandate from the public is a fallacy. We don't operate in an American system where you elect a president. You vote for party candidates and the largest party forms a government.

> [In response to an argument that compares voting to purchasing airline tickets.] Logical fallacy. Before you can vote you have to register to vote. Your registration is checked against public records. So voter registrations are validated. Purchasing airline tickets can be done by anyone with no validation at all. See the difference?

> Really weak straw man fallacy. I've never heard anyone claim the Constitution was infallible. Further, the fact that the amendment process was included means that the guys who wrote it didn't think it was infallible either!

> [In response to an argument that compares cars to houses.] That's a fallacy, there are more cars than people. There are far less available and affordable homes than there are people that can actually afford it. Building more homes don't lower prices because the prices aren't based off of supply and demand, but based off of the property value.

Nevertheless, the fact that almost half of the accusations were unsubstantiated is worrisome. As was explained in the previous section, this is a point on which argumentation theorists agree: accusations of fallacies must be substantiated for them to be reasonable.

Furthermore, as I was analyzing the tweets, I suspected that whether a fallacy was identified or not (criterion 1) influenced whether the accusation was substantiated (criterion 3). As the previous examples apparently show, when the fallacy was *not* identified, it was more likely that the accusation *was* justified. This would make sense, since the fallacy labels might be taken by ordinary arguers as substitutes for explanations or justifications. In order to test this hypothesis, let us see the percentage of accusations

that were substantiated depending on whether the fallacy was identified or not (see Fig. 3).
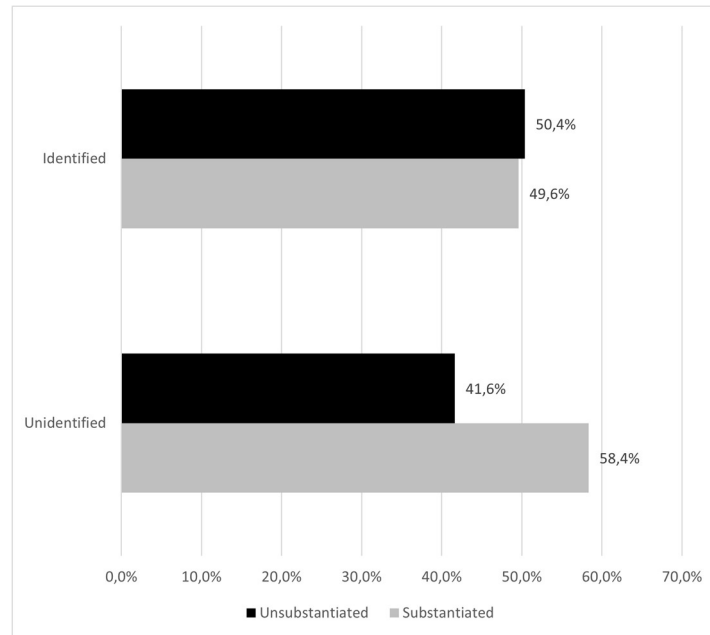

Figure 3. Identification and substantiation

We can see here that, whereas only 49,6% of the accusations that identified the fallacy were substantiated, the percentage of substantiated accusations rises up to 58,4% when the fallacy was *not* identified. An independent sample t-test showed that identified fallacy accusations were less likely to be substantiated (M = 0.496, SD = 0.5) than unidentified fallacy accusations (M = 0.584, SD = 0.494), and this difference was significant in our sample (t(863) = -2.54, p = 0.011). If this finding is confirmed by further studies, it could mean that teaching fallacy labels is *detrimental* to argumentative practice.

However, a possible caveat must be mentioned here. Even though, in principle, reasonable accusations of fallacies should include some kind of justification, some of the results might make us reconsider whether that is always the case[7]. It might be argued that, in certain dialogues, some labels are informative enough and a substantiation is not necessary. Perhaps this can be seen most clearly in some accusations of having committed an *ad hominem* fallacy. For example, the following dialogue was found in the sample:

A: We actually found the dumbest person on twitter. He truly is running his own race here.

B: Another argumentum ad hominem fallacy from you.

In cases like this one, it would seem that no substantiation is necessary, since it is obvious why B believes that A's comment is an *ad hominem* fallacy--although not every insult in a discussion is a fallacy, given that sometimes it is not even an argument, but I will leave this complication aside. If that is correct, then perhaps the requirement that accusations of fallacies must be justified might be subject to certain exceptions. In my view, however, simply saying that an argument fits into the pattern of an *ad hominem* does not entail that such an argument is a fallacy, since there are reasonable *ad hominem* arguments.

---

[7] I thank an anonymous reviewer for pointing out this possibility.

Consider: "John is very stupid and has a wild imagination, so I do not think that he has refuted the general theory of relativity as he says." Hence, I am inclined to think that simply saying "that is an *ad hominem*" is not informative enough--some kind of explanation is necessary. Even if, in certain specific cases, the label together with the context make it clear *to us* why there is a fallacy there, I would say that an explanation is still needed so that *the arguer* understands her fallacy--the fact that she committed the fallacy means that this is not clear to her.

Let us move on to criterion 4, about which we can perhaps be more optimistic: whether the accusations made any reference to the context in which the argument was put forward. Even though, in the global analysis presented in Fig. 2, 58,3% of the accusations made *no* reference to the context, we must bear in mind that that number includes the accusations that were *not* substantiated--and which therefore obviously did not make any reference to context either. If, however, we take into account only the substantiated accusations, the numbers are very different (see Fig. 4).
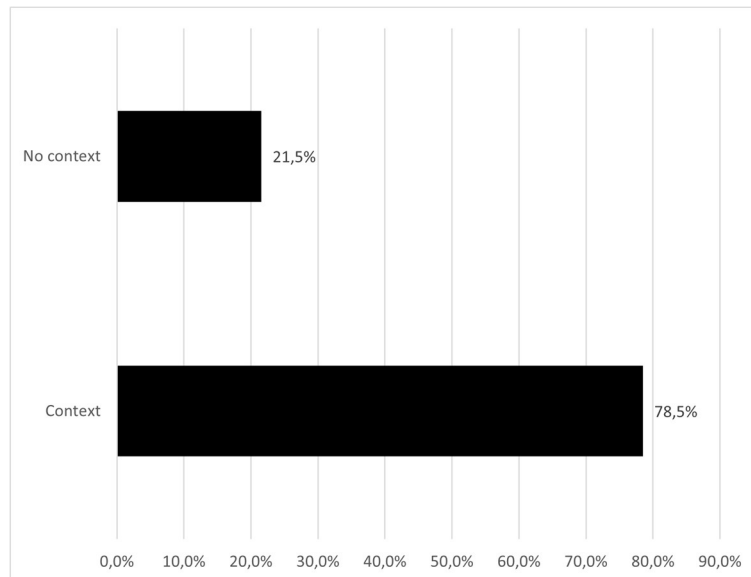

Figure 4. Context in substantiated accusations

Of those accusations that were substantiated, only 21,5% made no reference to the context, whereas 78,5% did. The following are some examples of substantiations without reference to the context (notice that they give us no information about the argument that is being criticized):

> The argument to moderation, the belief that the compromise position is always the most virtuous, is always a fallacy. It's never based on information, morals or ethics. Instead it's a mindless assumption that a midpoint is always right.

> Appeal to how many believe something is a logical fallacy. Nearly everyone thought the world was flat before. Knowledge is not a democracy.

> Nice try. Red Herring fallacy. A Red Herring argument is one that changes the subject, distracting the audience from the real issue to focus on something else where the speaker feels more comfortable and confident.

Contrast those with the following remarkable examples of accusations that discuss the context or the content of the argument:

> The "independent state legislature" 'doctrine' takes a fragment of the Constitution out of context, exaggerates the omission of an understood and implied element (the courts), and then prioritizes it over everything else the Constitution says. A massive "cherry-picking" fallacy.

> That's a logical fallacy. Going to the black market isn't as easy as going to Walmart. Many won't put in the extra work to get guns that way, many won't do it since it's now illegal and those that do get it you automatically know they are illegal so can take police action.

> Companies move to states that are less expensive to operate in all of time. This is an economic decision. The fallacy is you are taking this purely economic decision made by some large companies as their endorsement of the MAGA [Make American Great Again] way or an indictment of blue state leadership.

Criterion 5 (taxonomic technique) was assigned an affirmative value mainly when the accuser explained the scheme of the fallacy without getting into the content or the context of the argument. Given that, as we have just seen, references to context in substantiated accusations are relatively high, signals of taxonomic technique were obviously low (13,5%). The following are some examples of accusations that relied on the taxonomic technique:

> Haha! Fallacy of thinking. Slippery slope. Arguing that if an opponent were to accept some claim C1, then they would have to accept some other closely related claim C2, eventually leading to the conclusion that the opponent is committed to something absurd or obviously unacceptable.

> And your attack is *ad hominem*. Meaning you question my credibility, but not my arguments with evidence. It's a fallacy.

> *Tu quoque* is a type of *ad hominem* argument in which an accused person turns an allegation back on his or her accuser, thus creating a logical fallacy.

> You are asking the formal equivalent of, "If N+1 grains of sand can make a heap, then is it not the case that N grains of sand must also make a heap?" Answer is, No. To presume so is to fall for the fallacy, or the sorites paradox.

There were even less signs of accusers that relied on a problematic theory (5,3%). Some of the instances were due to a reliance on the taxonomic technique--which I take to be problematic--but there were also cases of questionable epistemological beliefs. The following examples show an exaggerated distrust of authorities, with the implication that no argument from authority can be reasonable:

> No shit and I applaud these doctors but that has been obvious for a long time. Trust your own instincts, no need for an appeal to authority fallacy to be convinced.

> "Experts say" = What follows will be a lie, but hey, it fits the narrative. Appeal to authority is in fact a logical fallacy--especially when, like with masks, the vax, lockdowns... The Science changes every month.

There were also cases of fallacies that were misconceived or poorly explained:

> Pointing out logical fallacies automatically invalidates your position due to it being a fallacy fallacy.

> No need to bring in defeated topics which are shown to contain false premises. That is the definition of appeal to tradition logical fallacy.

The relatively high percentage of "Dubious" results (9,6%) in this criterion is due to the appearance of several fallacies that are not included in traditional or current taxonomies, but that I was reluctant to consider problematic because they did not clearly clash with those taxonomies. Up to 41 categories of fallacies were of that kind. Some examples are "absolutism," "appeal to morality," "appeal to the future," "argument from incredulity," "bold face lie," "common sense," "God of the gaps," and "good guy with a gun." The following is a sample of such accusations:

> "Good guys with a gun" is a fallacy. Gun restrictions would be a better way to protect your family, you, and others. Try fighting for the right to life instead of the right to guns.

> 'Patriarchy' is a logical fallacy named golden hammer.

> This is just such a great example of the mind-projection fallacy. Without any irony, this dude assumes that the class discussing LGBT issues is what causes people to decide/declare/define themselves as nonbinary and bisexual.

> Once again: Isolation limits transmission, even if the isolation is imperfect & not up to the standards it should be. Hence why most infections were before or around quarantine. It's not my fault you abuse the nirvana fallacy to pretend otherwise.

Finally, for a more significant result, only 52% of the accusers were open to discuss their accusations, while the rest showed an unwillingness to do so. When an accusation was reasonably substantiated, the accuser was considered as open to discussion. But there were also more manifest signs of openness, such as asking questions:

> Then where do you get your definition? Just the "everyone knows" fallacy?

> "Perverting the nature of…" Why is perverting the nature of sex immoral? This is a classic is-ought fallacy. 1) Sex *is* for procreation therefore it *ought* to be used for procreation. This is exactly what Hume argued against and the primary problem with natural law.

> Racism is literally the ultimate expression of conservatism. Seeking to conserve traditional orders and hierarchies. They claim to be conservative themselves, are you saying they're all lying? Or is this a no true Scotsman fallacy you're attempting?

Also, expressions of uncertainty within the accusation:

> But I think the logical fallacy that he used was that causation equals correlation, which it does not.

> I'm wondering if there's an ecological fallacy here--at the group level I think you're right, but at the individual level cancellation will make a big difference. I'm not sure this is a case against debt cancellation but the point is very well taken.

And sometimes the accuser's openness was inferred from his or her attitude thorough the thread (i.e., the conversation), which I examined for each tweet. Unwillingness to discuss the issue, on the other hand, was sometimes shown in explicit statements that the conversation was over, such as in the following examples:

> We're not talking about fucking Joe Biden's failures and accomplishments as a legislator. That's a Red Herring and a Strawman fallacy in one. You suck at this. Funny part is you think you're a skilled logician. All done now, bye.

> I hope people will realize one day that there is a huge logical fallacy with the argument that life begins after breath. If so, when did life begin on Earth? Sorry, in

advance: but the truth is I most probably won't reply & it's up to you to work it out! Just think about it!

1st- comparing a random citizen wearing a sports hat is such a fallacy to comparing virtue signalling by forcing a Ukrainian flag in your name without any real action that it's not worth continuing discussing. You're a phony, clear and simple, and hope you one day do more than [Ukrainian flag].

Of course, although almost half of the accusers (48%) displayed an unwillingness to discuss their accusations, we do not know whether this has to do with their use of the concept of fallacy. They were, after all, engaged in heated discussions on contested issues--gun regulation, abortion, religion, transgender rights, vaccines, and even the existence of God! Moreover, hostility and unwillingness to listen to the other side seem to be unfortunately common on Twitter.

In this study, I have used seven criteria that, as I explained in the previous section, refer to characteristics that have been proposed in the literature as belonging to reasonable accusations of fallacies. It is, however, possible to identify a subset of those criteria as characterizing the minimum requirements that any accusation of fallacy must fulfill.[8] Arguably, if the fallacy approach to argument evaluation is to be pedagogically useful, at the very least accusations of fallacies must correctly identify the alleged fallacy and justify the accusation.[9] Let us then call a "correct" accusation one that gets positive answers for criteria 1 and 3 (it is identified and substantiated) and a negative answer for criterion 2 (it is *not* a misidentification). Fig. 5 shows the percentage of accusations that were correct in this sense.
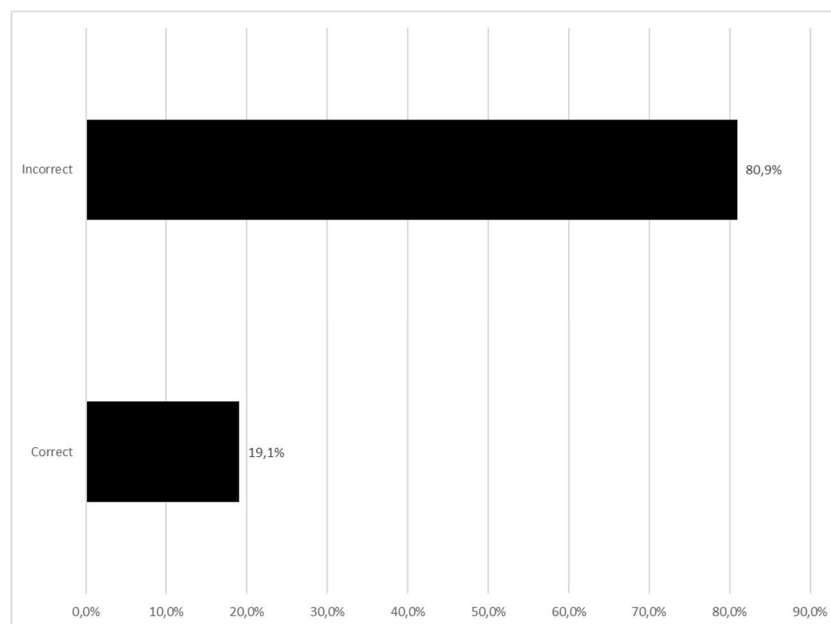

Figure 5. Correct accusations

---

[8] I am very grateful to an anonymous reviewer for suggesting this.

[9] Even though the results discussed above have cast doubt on the usefulness of identifying (i.e. labeling) fallacies, here I will leave that aside for a moment in order to adopt the perspective of fallacy theory. Arguably, from that perspective, identifying the specific fallacy is the right thing to do when accusing an arguer of having committed one.

Hence, less than 20% of the accusations were found to be correct. Although we lack enough data to draw any conclusion about why that is the case, we may speculate about some possible causes. A first obvious possibility is that people do not have deep enough knowledge of fallacy theory, relying instead on taxonomies that are presented on websites such as Wikipedia and that only provide brief and superficial explanations and do not emphasize the importance of justifying the accusations. If that is the case, the popularization of fallacies might turn out to be a problem. Another influencing factor may be the possibility, supported by the present study, that the identification (or labeling) of the fallacy discourages the justification of the accusation--so that it is less likely that both identification and justification will be present at the same time. And, finally, I believe that we should not rule out the possibility that, as some authors have argued, the fallacy approach to argument evaluation is not the right approach to teach people how to argue well. Further research should shed some light on this issue.

3.2. Kinds of fallacies

What fallacy labels appeared in the accusations that were analyzed? Even though examining this issue was not the main goal of the study (and therefore this section will be accordingly briefer) it might be of some interest to have a look at them. 512 accusations (52%) identified one or more kinds of fallacies. In total, the number of different categories of fallacies that were mentioned in the accusations was 107. Fig. 6 shows the percentages of appearance of the most frequent fallacy labels (those that appeared at least 10 times).
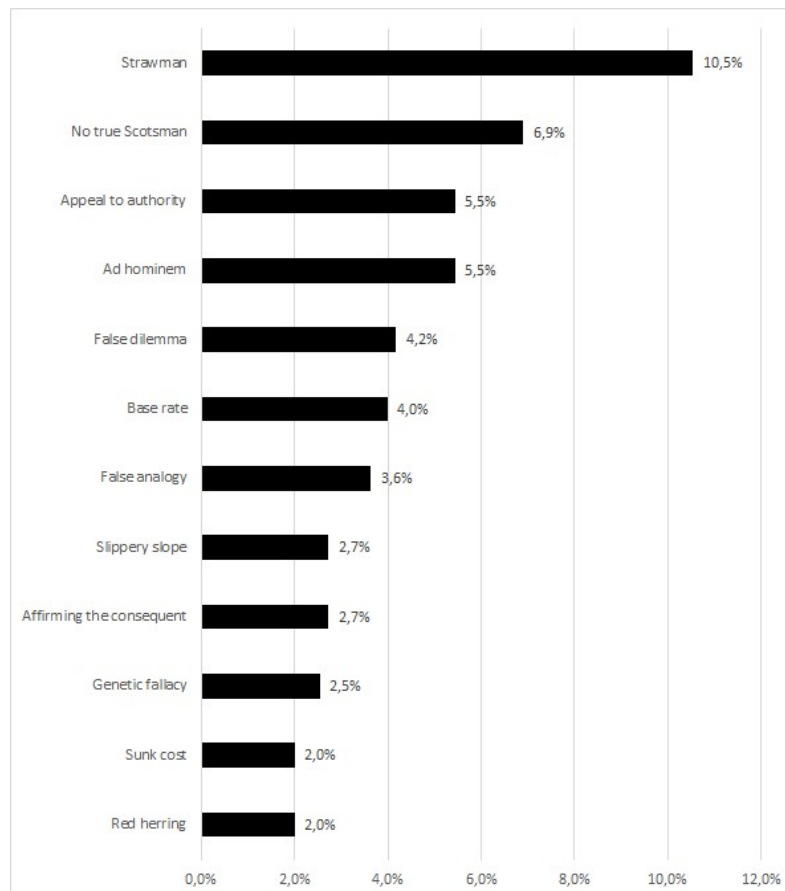

Figure 6. Most frequent fallacies

Strawman seems to be the most popular fallacy (10,5%), which might indicate that (not surprisingly) it is one of the most frequently committed fallacies or that it is one of the better-known categories of fallacies (or both). The no true Scotsman fallacy, first proposed by Antony Flew (1975: 47), is also one of the most popular and occupies here the second place (6,9%), even though it is rarely discussed in argumentation theory. It is followed by appeal to authority (5,5%), *ad hominem* (5,5%), and false dilemma (4,2%).

There are two fallacies in Fig. 6 whose presence some might find surprising. The first one is base rate (4%), which is a statistic error that arises when someone establishes a correlation between two phenomena without taking into account the general prevalence of one of them. It is, therefore, what we might call a specialized or scientific fallacy. The reason why this fallacy was so frequently mentioned in this study is that it has become popularized during the current discussions around COVID-19 and vaccines. Those who oppose vaccines sometimes point out that the majority of COVID-19 cases occur among vaccinated people, and therefore vaccines are ineffective. But, as the scientific community and many of the accusations in this study have insisted, such reasoning ignores the percentage of vaccinated people within a population--thereby committing the base rate fallacy. If, for example, 97% of the population is vaccinated, then a percentage of 57% vaccinated COVID-19 cases does *not* show that vaccines are ineffective--it shows precisely the opposite. Given that this is a specialized fallacy that is derived from Bayes' theorem, people who mention it might have more in-depth knowledge of fallacies as inferential mistakes.[10]

The second fallacy that I would like to discuss here is affirming the consequent (2,7%). Given that it is a formal fallacy, whose commission depends on avowedly having tried to put forward a deductively valid argument, its presence in this top-15 list is quite unexpected. One would think that people usually do not argue using propositional logic. However, it seems that they sometimes do, and here I came across one of those rare instances. As it turns out, *all* the accusations of affirming the consequent were made against *the same* argument, which it is worth quoting here:

> Your daily reminder that if 'transwomen are women' was true, the corollary mantra 'women are transwomen' would be equally true.

Strictly speaking, this is not an instance of affirming the consequent (if P then Q, Q, therefore P), but rather of what other users called "illicit conversion" or "illicit commutativity" (it P then Q, therefore if Q then P). But it comes close enough. The fallacy of illicit conversion or commutativity appeared only in 6 accusations and therefore did not make it to our list of frequent fallacies.

To conclude, some last comments about the other less frequent fallacies that do not appear in Fig. 6. In the first place, the fallacy of composition was mentioned in 8 accusations, but in *none* of them was the label used correctly. It was used twice to refer to illicit conversion (in reply to the aforementioned argument) and, most often, to hasty generalization. In no other category of fallacies do we find 100% misidentifications. So there is reason to believe that composition might be one of the most poorly understood fallacies.

---

[10] I thank an anonymous reviewer for suggesting this.

Secondly, it seems that new fallacies have been created that pertain only to a certain topic--quite independently of the argumentation scholarship and even of popular textbooks on fallacies. In this study, I found 6 instances of the so-called "no-limits fallacy," which does not appear on any argumentation textbook nor even on any popular webpage about fallacies that I know of. This fallacy is most clearly explained on a webpage about comic characters, where the fallacy is said to arise "when someone states that because something has not demonstrated any limits (or only certain limits) then it has none (or only the ones demonstrated)."[11] Apparently, it can be committed only in discussions about fights between comic characters--or at least I have not seen it in any other context. Since the fallacy is properly defined in its domain of application (manga comics), I have not considered that the accusations in which it appears rely on a problematic theory (criterion 6).

4. Conclusion

As I said, this study is but a first exploration on how fallacies are used in accusations in real discussions. Hopefully, more research will be conducted that will provide good grounds for determining how the accusations are usually made and how knowledge of the fallacies influences argumentative practice (for good or bad). Given that the concept of fallacy has been subjected to serious criticism and that many of the disputed claims are empirical (for instance, whether fallacies are common or whether learning about them improves the argumentative skills), I believe it is urgent to do this kind of research.

Some of the criteria that were chosen for the analysis turned out to be more informative than others. Some of them may even be questionable--when is an underlying theory of fallacy "problematic"? Nevertheless, even if certain results of this study should be disregarded, the relative independence of the criteria allows the rest of the findings to be of some value. In particular, it was striking to find that the percentage of accusations that were substantiated was significantly higher when the fallacy was not identified than when it was. Further research should look into this issue. If the use of fallacy labels tends to lead to making unsubstantiated accusations, then arguably we are doing a disservice by teaching them.

Finally, it is also interesting to note that taxonomies of fallacies seem to have taken a life of their own. Although the concept was created and spread by philosophers, ordinary arguers have developed it according to their own interests. 41 of the 107 different fallacies found in this study were not the kind of fallacies that one usually finds in the literature on argumentation--or even in the sciences. One of them ("no limits") was created by comic readers for a particular domain. Even though the most frequent fallacies are still mainly those considered and discussed in argumentation theory, the development of taxonomies of fallacies seems to be largely out of our hands now.

Acknowledgments

---

[11] https://character-level.fandom.com/wiki/No_Limits_Fallacy#16._No_Limits_Fallacy_(NLF)

the statistics and enlightened me with his knowledge of discussions about manga characters. A previous version of this paper was presented at the 4th European Conference on Argumentation (ECA) in Rome, 28-30 September 2022. I am grateful to Diane Liberatore, Hubert Marraud, Alexandru Cârlan, Michel Dufour, Michael Baumtrog, Dale Hample, Jennifer Schumann and Rahmi Oruç for their useful comments and their orientation for further research. Finally, I must also thank two anonymous reviewers, whose observations greatly improved the paper—even the title!

References

Blair, J. Anthony. 1995. "The place of teaching informal fallacies in teaching reasoning skills or critical thinking." In *Fallacies: Classical and contemporary readings*, ed. by Hans V. Hansen, and Robert C. Pinto, 328–338. University Park, PA: The Pennsylvania State University Press.

Blair, J. Anthony. 2023. "Teaching the fallacies." *Argumentation* 37(2): 247–251.

Boudry, Maarten, Fabio Paglieri, and Massimo Pigliucci. 2015. "The fake, the flimsy, and the fallacious: Demarcating arguments in real life." *Argumentation* 29: 431–456.

Eemeren, Frans H. van, Bart Garssen, and Bert Meuffels. 2009. *Fallacies and judgments of reasonableness*. Dordrecht: Springer.

Finocchiaro, Maurice A. 1981. "Fallacies and the evaluation of reasoning." *American Philosophical Quarterly* 18: 13–22.

Finocchiaro, Maurice A. 1987. "Six types of fallaciousness: Toward a realistic theory of logical criticism." *Argumentation* 1: 263–282.

Finocchiaro, Maurice A. 2007. "Arguments, meta-arguments, and metadialogues: A reconstruction of Krabbe, Govier, and Woods." *Argumentation* 21: 253–268.

Flew, Antony. 1975. *Thinking about thinking*. London: Fontana.

Hitchcock, David. 1995. "Do the fallacies have a place in the teaching of reasoning skills or critical thinking?" In *Fallacies: Classical and contemporary readings*, ed. by Hans V. Hansen, and Robert C. Pinto, 319–327. University Park, PA: The Pennsylvania State University Press.

Hundleby, Catherine. 2010. "The authority of the fallacies approach to argument evaluation." *Informal Logic* 30: 279–308.

Jason, Gary. 1987. "Are fallacies common? A look at two debates." *Informal Logic* 8: 81–92.

Johnson, Ralph H., and J. Anthony Blair. 1994. *Logical self-defense*. New York: McGraw-Hill.

Massey, Gerald J. 1975. "Are there any good arguments that bad arguments are bad?" *Philosophy in Context* 4: 61–77.

Massey, Gerald J. 1981. "The fallacy behind fallacies." *Midwest Studies in Philosophy* 6: 489–500.

Tindale, Christopher W. 2007. *Fallacies and argument appraisal*. New York: Cambridge University Press.

Toulmin, Stephen, Richard Rieke, and Allan Janik. 1984. *An introduction to reasoning*. 2nd ed. New York: Macmillan.

Walton, Douglas N. 1995. *A pragmatic theory of fallacy*. Tuscaloosa: University of Alabama Press.

Willingham, Daniel T. 2008. "Critical thinking: Why is it so hard to teach?" *Arts Education Policy Review* 109: 21–32.

Author's address

José Ángel Gascón

Departamento de Filosofía

Facultad de Filosofía

Universidad de Murcia

Campus de Espinardo

30100, Murcia

Spain

jgascon@um.es