languages

MDPI

Article

Do People Perceive the Disagreement in Straw Man Fallacies? An Experimental Investigation

Jennifer Schumann

English Department, University of Fribourg, 1700 Fribourg, Switzerland; jennifer.schumann@unifr.ch

Abstract: So far, experimental studies on the straw man have targeted the misrepresentational dimension of this fallacy. In order to provide a more detailed understanding of the way the straw man is perceived, the focus of this paper lies on the refutational dimension. In two experiments, I will assess (1) if people are sensitive to the underlying disagreement expressed through the use of a straw man and (2) if question wording plays a role for the perception of disagreement. The results of the experiment show that participants indeed notice easily that the person performing a straw man disagrees with his opponent. It also emerges from the experiment that the difference between a positive or negative formulation of the experimental questions does not affect the perception of disagreement in the straw man. The underlying disagreement in the straw man is thus perceived either way.

Keywords: straw man; disagreement; wording effect; fallacies; empirical validation

check for updates

Citation: Schumann, Jennifer. 2022. Do People Perceive the Disagreement in Straw Man Fallacies? An Experimental Investigation. *Languages* 7: 111. https://doi.org/10.3390/languages7020111

Academic Editors: Steve Oswald, Juana M. Liceras and Raquel Fernández Fuertes

Received: 15 January 2022
Accepted: 27 April 2022
Published: 3 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The straw man fallacy is commonly defined on the grounds of two fundamental factors: the misrepresentational and the refutational aspect (e.g., Aikin and Casey 2016; Oswald and Lewiński 2014). Whereas the misrepresentational dimension of the straw man is mainly oriented towards the form, the refutational dimension is more concerned with the function (e.g., Lewiński 2011). For there to be a straw man, two conditions have to be met. It is necessary that some type of misrepresentation of the opponent's position takes place. But this misrepresentation also has to fulfil a specific goal, which is to refute the views of the opponent. The distortion of the content thus serves as the basis of the attack that is carried out on an adversary position. Since the aim of the straw man is to refute the opponent's position, it becomes clear that the person uttering the straw man does not agree with the person they attack. A straw man fallacy thus expresses some sort of disagreement between two parties.

The straw man has mainly been investigated from perspectives adopting a theoretical angle (see e.g., Aikin and Casey 2016; de Saussure 2018; Lewiński 2011; Oswald and Lewiński 2014), but empirical approaches to the topic are increasing (Bizer et al. 2009; Schumann et al. 2019, 2021; Schumann 2022). In this paper, I address the straw man from an experimental point of view, and the aim is twofold. First, I provide an analysis of two factors in relation to the refutational dimension of the fallacy. More specifically, I investigate the perception of disagreement in straw men on the one hand and the wording effect in relation to the questions asked in the experimental design on the other hand. Second, I show how research on fallacies can benefit from more empirically oriented studies as presented in this paper.

This paper is structured as follows: Section 2 discusses the notion of disagreement in relation to argumentation in general and the straw man fallacy more specifically. Section 3 elaborates on the importance of positive and negative formulations of experimental questions and their potential impact on the way the disagreement is perceived. The experiment

investigating the perceived refutational dimension of the straw man fallacy will be presented in Section 4. After a short introduction to the experiment, a detailed presentation of the participants, the materials, the procedure, the analysis, and the results will follow, as well as a discussion of the results. I will conclude this paper with some general remarks in Section 5.

2. Disagreement in Argumentation and the Straw Man

Disagreements can be found everywhere. People disagree over political opinions, over scientific studies, and over things that are as casual as what to eat for lunch on a Monday. A disagreement can take place in written and oral form, in private and public space, in person and online, in large and small groups, and so on. When arguing, people's views and opinions diverge; they realize that part of the position they hold is incompatible with the suggestions uttered by the opposing party. In pragma-dialectical terms, argumentation is about people having a conflict of opinion, a disagreement that the parties involved in the argumentation seek to resolve in a critical discussion (e.g., van Eemeren and Grootendorst 1992, 2004). On the grounds of a disagreement, the arguers thus proceed to an argumentative exchange, acting on and reacting to the opponent's respective position and arguments. The disagreement is therefore crucial for the establishment and process of an argumentation. During a critical discussion, the argumentative exchange passes through four different stages, namely the confrontation stage, the opening stage, the argumentation stage, and the concluding stage (van Eemeren et al. 2002; van Eemeren and Grootendorst 2004). Each stage entails different functions. In the first stage, i.e., the confrontation stage, the disagreement between the parties becomes apparent. This is where the parties realize that they have conflicting opinions. The expression of a disagreement is, however, not limited to the first stage of the critical discussion. According to Tseronis (2021), a disagreement is considered as an argumentative move that can take different forms and can occur in all stages of the critical discussion. A disagreement is initially noticed in the confrontation stage, then arguers share their disagreement with the starting points in the opening stage, exchange arguments which can be challenged and disagreed on in the argumentation stage, and they might still disagree on the outcome of the argumentative exchange in the concluding stage. As the pragma-dialectical model (see e.g., van Eemeren and Grootendorst 2004) and Tseronis (2021) show, disagreement is at the core of argumentation. However, there are instances of argumentation where an argumentative exchange arises for other reasons. Sometimes, people are more concerned with maintaining an agreement instead of resolving a disagreement. In these situations, the expressed arguments reinforce shared views. In regard to the straw man, Talisse and Aikin (2006) and Aikin and Casey (2022) argue that the straw man is a particularly effective argumentative move when the audience already shares the impression that the target of the straw man attack is weak. But as the works of van Eemeren and Grootendorst (2004) and Tseronis (2021) also illustrate, the presence of a disagreement does not necessarily always mean that the argumentative exchange is unreasonable. On the contrary: in an ideal case, an argumentation emerging from disagreement follows principles of reasonableness. In this line of thought, the ideal model for critical discussion provided by the pragma-dialectical framework describes an approach where normative principles such as the rules for critical discussion are required for the argumentation to remain reasonable. This entails that any violation of the principles leads to fallacious argumentative moves. In real argumentative settings, arguers seek to defend their point of view, and when the willingness to win the argumentative exchange prevails, the risk of being unreasonable increases. When arguing, the people involved always balance on a fine line between remaining reasonable and being effective. This balancing act is described as strategic maneuvering in the extended version of the pragma-dialectical framework (van Eemeren and Houtlosser 1999). Arguers derail into fallacious argumentation when they are more concerned with the rhetorical effectiveness of what they say, rather than remaining within the scope of reasonable argumentation. This goes to show that a disagreement is not unreasonable in itself, but the argumentative exchange that emerges from it can be, in

cases where the arguers are driven by rhetorical effectiveness. One example for such an occurrence of unreasonable argumentation is the use of fallacies to bring across a point.

Through the use of a fallacy such as the straw man, the already existing disagreement between the parties can be highlighted or reinforced. The misrepresentation of the opponent's position serves as a basis for the refutation of said position, showing that the attacking party is in disagreement with the target of the original position. From this point of view, the refutational dimension of the straw man fallacy expresses an attack directed towards the opponent. Much like disagreement in general, the straw man fallacy is not linked to a specific stage of the critical discussion. This fallacy does not have a specific form or structure that makes it difficult to perceive it as an argument scheme, contrary to other fallacious arguments, e.g., the appeal to authority (see e.g., Walton et al. 2008). The reason behind this lies within the fact that the straw man is a fallacy operating on somebody else's argument, rather than being an argument on its own (see also Oswald and Lewiński 2014). This means that arguers can perform a straw man at any given time during the critical discussion. As a consequence, any point of view or argument uttered by the opponent can be refuted at any stage of the argumentative exchange. Seen from a pragma-dialectical perspective, fallacies hinder the continuation of a reasonable exchange and ultimately the resolution of a conflict of opinion (van Eemeren and Grootendorst 1992). When a speaker performs a straw man, they distort the original argument uttered by their opponent and use this distorted version as an attack, discrediting them and their argument in order to refute the initial position. Through this refutation, the speaker of the straw man puts the target into a position of weakness. In fact, as de Saussure describes (de Saussure 2018), they present themselves as having superior rhetorical competences compared to their opponent. In doing so, the person uttering a straw man creates an obstruction because the opponent now has to justify how the speaker of the fallacy has misrepresented their position. As a consequence, performing a fallacy on someone hinders the continuation of a reasonable critical discussion aimed at resolving the disagreement. In fact, through the refutational dimension of the straw man, the disagreement between the parties involved becomes even more apparent. The attacking person, i.e., the speaker of the straw man, shows that they do not endorse the same position by refuting what has been said initially. In other words, a disagreement between two parties can be highlighted through the use of a straw man fallacy because the refutational dimension makes the attack more explicit and increases the visibility of the arguers' conflicting opinions.

Using a straw man constitutes one of many possibilities to signal disagreeing positions in argumentative discourse. Generally, when thinking about disagreement, one of the first characteristics that comes into mind is the presence of diverging opinions. But even when people disagree and endorse two different positions, they do so with a sufficient number of elements that both interlocutors have in common (Aikin and Talisse 2020, chp. 8). In other words, two persons may have opposing views, but they can generally agree on the required terms for successfully resolving their conflict of opinion, e.g., to be reasonable, to remain truthful, to verify information that has been shared. They share the same background knowledge and principles. Consider the following scenario: Erin and Melissa talk about their favorite movie director, namely Quentin Tarantino. They disagree over the year in which Tarantino's recent movie, *Once upon a time in Hollywood*, was released. Erin defends the position that the COVID-19 pandemic had already started, whereas Melissa is convinced that the movie was in theatres the year before the pandemic. After exchanging arguments for a while and not reaching an agreement, they turn to the Internet Movie Database (© Imdb.com Inc., Seattle, WA, USA) in order to verify the date of release. Through the information provided by this platform, they are able to find out the precise release date and reach an agreement in the discussion. In this case, the disagreement can be resolved because both persons involved in the critical exchange turn to an independent source of information through which they are able to access the facts. Sometimes, however, the frontiers between both parties are more extreme and the disagreement rips a deep gap between the arguers and their positions. For Shields (2018), a

deep disagreement arises when two parties have a fundamentally different understanding of the same concept. This entails that the argumentative exchange around it is construed on different principles. Consequently, there is no apparent solution to the disagreement as a general consensus on the background of the concept (e.g., the exact release date of a movie, as mentioned above) is lacking. Aikin and Talisse (2020, chp. 8) approach the notion of deep disagreement from a similar perspective and consider that such a type of disagreement arises when the people involved in an argumentation do not only disagree on the arguments and opinions expressed in the argumentative exchange, but also on the factors that determine the disagreement. In these cases, it becomes difficult to find common ground, and the people arguing are often not willing to make concessions or to accept opinions diverging from their own. Yet, the depth of a disagreement has to be seen as a scalar notion, as Aikin (2019) explains. Some disagreements might be deeper than others. However, even during the deepest of disagreements, parties agree on one aspect, namely the fact that they disagree. Consequently, a disagreement cannot reach an absolute depth (Aikin 2019; Aikin and Talisse 2020). They can, however, reach depths where the arguers deviate from reasonable argumentation. The persons involved in the argumentative exchange are not always able to successfully defend their position, which can lead them to use arguments that are dialectically flawed and thus more oriented towards the pursuit of rhetorical gains and effectiveness. Fallacies can be a product emerging from these circumstances. Consequently, using fallacies to express a disagreement can be a manifestation of its depth. The deeper the disagreement becomes, the more difficult it gets to find reasonable arguments that are accepted by the opponent. For this reason, one might be tempted to commit a fallacious reasoning in order to get their views across. Fallacies such as the straw man can therefore arise from a context of deep disagreement and can be used to express and reinforce these types of diverging opinions. This fallacy is based on a misrepresentation of the opponent's position in order to more easily refute it (see e.g., Aikin and Casey 2016; Oswald and Lewiński 2014). Through the refutational dimension, the attacking party shows that they disagree with the position expressed by the opponent. The issue with fallacies such as the straw man is that they do not only express a disagreement, but they do so in an erroneous way. The fact that the straw man distorts the opponent's initial viewpoint in the argumentative exchange hinders a rational continuation of the discussion and hinders a resolution of the disagreement. It does so by opening room for a discussion on the principles of reasonable argumentation. The argumentative exchange is thus not exclusively focused on factual aspects anymore, but has the potential to deviate into a discussion about the elements required for sound reasoning and argumentation. It can become meta-argumentative, resulting in an argument about the argument (Finocchiaro 2011; see also Oswald and Lewiński 2014 for their discussion of the straw man's meta-discursive aspect). The problem with the use of fallacies such as the straw man in argumentative exchanges is that they are not a tool for resolving a difference of opinion. On the contrary: they might actually deepen the disagreement even further.

The perception of disagreement in the context of straw man fallacies has not been tested so far. Even though experiments such as those conducted by Schumann et al. (2019, 2021) have shown that participants easily spot fallacious statements, they never explicitly tested whether participants also detected the underlying disagreement between the interlocutors. Consequently, it remained unclear whether people were sensitive to the disagreement underlying the straw man. However, as the fallacious nature of the straw man was always perceived more negatively (see Schumann et al. 2019, 2021), it seems probable that the participants noticed the underlying disagreement between the person uttering the straw man and the target. To remove the room for speculation and to provide solid answers, an experiment aimed at the refutational characteristic of the straw man was in order. The aim of this paper is therefore to focus in-depth on the more function-oriented side of the straw man, showing that participants are aware of the disagreement between the interlocutors when straw men are used.

3. The Role of Question Wording for the Perceived Agreement

One aspect that must also be considered when conducting experimental studies is the way the experimental questions used in a survey are formulated. A number of studies (e.g., Kalton et al. 1978; Kamoen 2012; Rugg 1941; Rugg and Cantril 1942; Schuman and Presser 1977) have investigated the role of question wording in polls targeting public opinion. These studies were interested in the way the wording of the selected questions may influence the attitude people have towards a certain aspect. Questions can be formulated with a positive wording or a negative wording (e.g., *agree/disagree*, *allow/forbid*, *yes/no*, etc.). Studies on question wording assess if and how the type of formulation affects the attitudes reflected in the participants' answering patterns. The results of these studies have been mixed: whereas some studies have found an effect (e.g., Rugg 1941), showing that negative question wording leads to stronger impacts on attitudes, others have not found any effect at all (e.g., Bishop et al. 1988).

In recent studies, Kamoen (2012) and Kamoen et al. (2011) have provided extensive work in order to understand the cognitive underpinnings of the question wording effect. They assessed the processes that conduct to different answering patterns for positively and negatively formulated questions. In an eye-tracking experiment, Kamoen et al. (2011) investigated the cognitive processes that are activated when answering attitude questions by assessing at which stage of the process—the comprehension-retrieval stage (understand and combine information) or the mapping stage (selecting an answer)—the wording effect takes place, as well as the role of different words pairs and scale size on the effect. To comprehend how question wording influences the responses given by a person, it is important to understand how the processes of the first stage impact the formation of a judgment and consequently the answering choice in the second stage. The aim was therefore to assess whether the wording effect takes place during the first or the second stage. The researchers designed two versions of the same experiment working with 4 question-clusters that included either a 2-point scale with the options *yes* or *no* or a 5-point scale with *agree* and *disagree* at both extremes, as well as positive and negative wording of different words pairs (Kamoen et al. 2011). This enabled them to test for the influence of scale size and word pairs on the attitudes. The results showed that the wording effect had an impact on both versions, but they also highlighted that the effect is much stronger in the case of 2-point scales and less consistent for 5-point scales (Kamoen et al. 2011, pp. 369–75). In respect to the processing time, the researchers did not find any significant difference between the reading times either for positive and negative wording in the questions or for the answers on either scale (Kamoen et al. 2011, pp. 369–75). However, Kamoen et al. (2011, pp. 369–75) found that people were more likely to reread a question that was formulated in a negative manner, which led to the conclusion that answering was more difficult on negative questions, leading to an increased difficulty in the mapping stage. Overall, the eye-tracking experiment by Kamoen et al. (2011) showed that the wording effect is not unequivocally proven. There are some aspects that point towards its existence, notably in the case of 2-point scales, but the results cannot be generalized. The data have shown that there are no processing differences when it comes to the comprehension-retrieval stage as reading times were comparable, but the effect seems to be more associated with the mapping stage.

Question wording has been shown to be a factor contributing to the perception and evaluation of content in surveys for the most part. So far, the experimental studies on the straw man (see Bizer et al. 2009; Schumann et al. 2019, 2021) have not assessed the role of question wording in relation to the perception of the fallacy. As the effect cannot be generalized, I decided to investigate how question wording affects the perception of disagreement in straw men. This paper focuses on the perceived disagreement in straw man fallacies, i.e., the people's ability to spot that the speaker of the straw man does not endorse the same opinion as their target. Asking about the perceived disagreement represents a negative formulation. To assess a possible wording effect, the results needed to be compared to a more positive formulation targeting the perceived agreement.

4. Experiments Testing the Perceived Refutational Dimension

This section presents two experiments that specifically targeted the refutational dimension of the straw man: one experiment investigated the perceived agreement between the speaker and the target of the straw man, and one experiment assessed the perceived disagreement between both interlocutors. The aim of the experiment was to show that participants are able to spot the disagreement between the interlocutors and to investigate whether the wording of the questions guides the evaluation of the perceived disagreement.[1] As the previous experiments of Schumann et al. (2019, 2021) have demonstrated, participants are sensitive to straw man fallacies, as they always attributed lower scores when asked about their degree of agreement with the person who uttered the misrepresentation. Given this result, I argue that the detection of fallacies may also be related to the people's competence in detecting an underlying disagreement between the interlocutors. The experiments described in this section investigate the hypothesis that participants detect the disagreement between the interlocutors, leading to lower scores for fallacious compared to non-fallacious items.

The second aspect investigated in the experiments is the role of question wording for the perception of disagreement. Using a positive or negative formulation for the questions may also potentially affect the perception of disagreement in the context of fallacious argumentation. If the wording effect obtains in the case of straw man fallacies, the questions targeting the perceived disagreement will lead to lower scores than the questions targeting the perceived agreement.

Taken together, these two dimensions will provide an empirically grounded insight into the refutational dimension of the straw man fallacy. It will demonstrate that people are not easily fooled, and that they are able to detect a disagreement between two interlocutors. In order to assess these elements, the following sub-sections will provide an overview of the experimental design and discuss the results obtained from the studies.

4.1. Participants

For the experiment targeting the perceived agreement, I recruited 37 French-speaking participants (14 female and 1 non-binary, mean age: 31, age range: 18–62) on the crowdsourcing platform Prolific© (Prolific, Oxford, UK). Each participant was rewarded £1.88 for their participation. At the beginning of the study, the participants were asked to give their informed consent. On average, participants completed the task in 20 min. For the experiment aiming at the perceived disagreement, 38 French-speaking participants (16 female, mean age: 30, age range: 18–55) were recruited via the crowdsourcing platform Prolific© (Prolific, Oxford, UK). All participants received £1.88 for participating in the study. Each participant was asked to give their informed consent before starting with the experimental task. Participants needed 21 min on average to complete the study.

4.2. Materials

The materials for both experiments consisted of the same 40 dialogues that Schumann et al. (2019, 2021) used for their previous studies. Each dialogue contained two turns: the first, illustrated in 1., was expressed by Barbara and remained unchanged through all experimental conditions.

1. Barbara: Il est crucial de mieux soutenir les jeunes parents parce qu'avoir un enfant signifie beaucoup de charges financières.
   Barbara: It is crucial to better support young parents because having a child means having a lot of financial responsibility.

Barbara's statement always introduced a standpoint in the first segment "It is crucial to better support young parents" and an argument in the second segment "having a child means having a lot of financial responsibility". The argument was linked to the standpoint via the causal connective *parce que* (*because* in English).

The second turn of the dialogue, illustrated in 2. to 5., was always expressed by a person named Alexandre and could appear in four different conditions. The first condition in 2. constitutes a case of straw man, where the fallacious argument is introduced with the causal connective *puisque* (closest to *since* in English).

2. Alexandre: Augmentons les allocations familiales puisqu'on ne pense qu'à l'argent.
   Alexandre: Let's raise the family allowance since it is only about the money.

In the first segment, "Let's raise the family allowance", Alexandre introduces a possible consequence of the argument provided by Barbara. This segment remained identical throughout all experimental conditions. The second segment, "it is only about the money", represents a distorted argument based on Barbara's initial argument. The second segment constitutes a fallacious reformulation, resulting in an exaggeration and thus amounting to a straw man. The same sentence was used for the second condition, which therefore also represents a case of straw man, but this time, as illustrated in 3., the fallacious argument was presented without a connective, indicating the causal coherence relation between the segments.

3. Alexandre: Augmentons les allocations familiales. On ne pense qu'à l'argent.
   Alexandre: Let's raise the family allowance. It is only about the money.

The third experimental condition, illustrated in 4., is a non-fallacious reformulation of the argument expressed by Barbara.

4. Alexandre: Augmentons les allocations familiales puisque les parents sont sous pression économique.
   Alexandre: Let's raise the family allowance since the parents are under economic pressure.

Again, the first segment remains unchanged. The critical element is located in the second segment, in this case a possible reformulation introduced with a connective explicitly marking the causal coherence relation. The fourth condition, illustrated in 5., is the same sentence as presented in 4., but the segments are juxtaposed instead of linked with a connective.

5. Alexandre: Augmentons les allocations familiales. Les parents sont sous pression économique.
   Alexandre: Let's raise the family allowance. The parents are under economic pressure.

The four experimental conditions were attributed to four different lists using a Latin square design, so that every participant only saw one of the conditions per item. The participants were asked to read 10 items per condition, amounting to 40 dialogues in total. The experiments on the perceived agreement and the perceived disagreement took place separately, meaning that participants only responded to questions targeting either the agreement or disagreement between interlocutors.

4.3. Procedure

The experiment was set up on the Qualtrics© (Qualtrics LLC, Provo, UT, USA) platform and distributed to the participants through a weblink on the crowdsourcing platform Prolific© (Prolific, Oxford, UK). At the beginning of the experiments, the participants received some general information on the study: they were told that were about to read forty dialogues of various contents between two persons named Barbara and Alexandre and that they would be asked four questions for every dialogue. The participants were instructed to carefully read the dialogues and to respond spontaneously to the questions, as there were no right or wrong answers. After giving their informed consent to participate in the study, some socio-demographic questions (gender, age, native language, place of residence) followed. To familiarize the participants with the experimental task, they were shown two exemplary dialogues. The explanations to the examples not only provided specification on how to read the dialogues, but also introduced the concepts of proposal and reason that were used to formulate the questions. From there, the participants moved

on to the experiment part, during which they read forty dialogues about various topics. The dialogues appeared in a randomized order, and the participants were asked to evaluate each of them on the basis of four questions. The answers were given on a 6-point Likert scale ranging from "non, pas du tout" (i.e., 'no absolutely not') to "oui, tout à fait" (i.e., 'yes, absolutely'). It was also possible to select an additional option, namely "je ne sais pas" (i.e., 'I don't know'), if the participants were not able or did not want to give an answer.

Whereas the items remained the same, the questions changed between both experiments. This choice was made to compare the effect of a more positive formulation of the questions, targeting the degree of agreement between Barbara and Alexandre, with the impact of a more negative formulation, targeting the degree of disagreement between both speakers. In the first experiment, the participants were asked about the perceived agreement between the interlocutors. The first question illustrated in 6. targeted the perceived agreement of Alexandre with the standpoint, i.e., the first segment of Barbara's initial statement.

6. Selon sa réponse, est-ce qu'Alexandre semble en accord avec la proposition de Barbara?
   Based on his response, does Alexandre seem to be in agreement with Barbara's proposal?

The second question presented in 7. aimed at the perceived agreement of Alexandre with Barbara's argument expressed in favor of the standpoint, i.e., the second segment of the initial standpoint.

7. Selon sa réponse, est-ce qu'Alexandre semble en accord avec la raison donnée par Barbara ?
   Based on his response, does Alexandre seem to be in agreement with the reason given by Barbara?

The results of the study presented in Schumann et al. (2019) illustrated that the degrees of acceptability for straw men based on a misrepresentation of the standpoint were lower compared to the acceptability of straw men construed on the basis of a distortion of the argument. In other words, the participants were more inclined to accept a misrepresentation of the argument, as it only reflected a doubt of the argument supporting a position and not the position in itself. The experimental design in the present paper takes this distinction into account by asking specifically for the perceived agreement between Alexandre and Barbara's standpoint, as shown in 6., or the argument, as shown in 7. In the second experiment, the participants were asked about the perceived disagreement between the interlocutors, Barbara and Alexandre. As can be seen from the first question illustrated in 8., the wording remained the same, with the exception that the word *agreement* was replaced with the word *disagreement*. Again, the first question targeted the perceived disagreement between Alexandre's and Barbara's point of view.

8. Selon sa réponse, est-ce qu'Alexandre semble en désaccord avec la proposition de Barbara?
   Based on his response, does Alexandre seem to be in disagreement with Barbara's proposal?

The question shown in 9. is the negative counterpart to the question presented in 8. Here, the focus lies on the perceived disagreement between Alexandre and Barbara's argument.

9. Selon sa réponse, est-ce qu'Alexandre semble en désaccord avec la raison donnée par Barbara?
   Based on his response, does Alexandre seem to be in disagreement with the reason given by Barbara?

As discussed above, it has been shown in Schumann et al. (2019) that a misrepresentation of the argument, which in this case was based on a less plausible rephrasing of the initial content, is better accepted than a misrepresentation based on a standpoint. In a first scenario, the perceived disagreement with the standpoint could lead to higher scores,

meaning that people have stronger attitudes towards the disagreement when it comes to their perception of the overall position. As a disagreement with the standpoint entails a higher rupture with the original position, it could result in more visibility and thus lead to a stronger impact on the attitudes expressed in relation to the perceived disagreement. In other words, the disagreement could become more visible. In the second, more probable scenario, people could react more strongly when the experimental questions aim at the argument. In this case, the overall position is not negated, and the disagreement focuses on the element linked to the initial position. The disagreement with the opponent's argument thus appears to be more tangible and less invasive rather than a disagreement with the opponent's standpoint. However, people are expected to detect the disagreement in straw men in either case, i.e., with the standpoint or the argument.

4.4. Analysis and Results Targeting the Perceived Agreement

The analysis was performed with a 2 × 2 × 2 mixed ANOVA with two within-subject factors, i.e., the type of argument (fallacious vs. non-fallacious) and the use of *puisque* (present vs. absent), as well as one between-subject factor (agreement vs. disagreement).[2] Only answers given on a scale from 1 to 6 were included in the data, as the additional option 'I don't know' was treated as missing data.

The scale used for the experiment on the perceived disagreement was the same as the one used for the experiment on the perceived agreement. As a consequence, the polarity of the highest and lowest point of the scale regarding the agreement was inverted compared to the same scale targeting the disagreement. In other words, if participants were convinced that Alexandre agrees with Barbara, they should respond 'yes, absolutely' to the question 'Based on his response, is Alexandre in agreement with Barbara's proposition', but because of the reverse value due to the negative wording, they should respond 'no, absolutely not' to the question 'Based on his response, is Alexandre in disagreement with Barbara's proposition'. For better comparison between the experiments, I thus converted the scores obtained in the experiment, targeting the perceived disagreement. I used ©*SPSS Statistics* (IBM Corp., Armonk, NY, USA), applying the rule shown in 10:

10. Converted Score = 7 − Old Score

To compute the new score, the old score is subtracted from a value that is 1 higher than the highest value of the scale (6-point scale, thus 7).

Table 1 contains the means and standard deviations for the experiment that targeted the perceived agreement and disagreement between Barbara and Alexandre.

Table 1. Means and SD of the perceived agreement/disagreement between the interlocutors.

	Agreement		**Disagreement**	
	Mean	**SD**	**Mean**	**SD**
Question 1: Proposition				
Fallacious argument with *puisque*	4.48	0.74	4.29	0.66
Fallacious argument without *puisque*	4.48	0.69	4.43	0.68
Non-fallacious argument with *puisque*	4.73	0.59	4.56	0.65
Non-fallacious argument without *puisque*	4.68	0.69	4.63	0.63
Question 2: Reason				
Fallacious argument with *puisque*	4.05	0.71	3.87	0.79
Fallacious argument without *puisque*	4.22	0.73	4.00	0.72
Non-fallacious argument with *puisque*	4.75	0.67	4.46	0.74
Non-fallacious argument without *puisque*	4.76	0.69	4.63	0.59

On the first question targeting the agreement between Alexandre and Barbara's proposal, the results showed a significant effect for the type of argument. Participants perceived higher agreement scores between the interlocutors for non-fallacious statements (M = 4.65) compared to fallacious statements (M = 4.42) [$F1$ (1, 74) = 27.41, $p < 0.001$]. For the second

factor, i.e., the role of the connective *puisque* as a marker of disagreement, no effect emerged [$F1$ (1, 74) < 1]. Additionally, the analysis did not return any significant effect for the between-subject factor, namely the question wording [$F1$ (1, 74) < 1]. Furthermore, the analyses did not reveal any type of interaction between any of the following factors: type of argument * wording [$F1$ (1, 74) < 1], connective * wording [$F1$ (1, 74) = 1.74, $p = 0.191$], type of argument * connective [$F1$ (1, 74) < 1], type of argument * connective * wording [$F1$ (1, 74) < 1].

For the question aiming at the agreement between Alexandre and Barbara's reason, the pattern is similar to the previous question. Participants attributed higher agreement scores in cases of non-fallacious statements (M = 4.65) compared to their fallacious counterparts (M = 4.04) [$F1$ (1, 74) = 110.79, $p < 0.001$]. The results of the connective *puisque* show that arguments were better accepted when they were simply juxtaposed to the previous segment (M = 4.4) instead of being introduced with a connective (M = 4.28) [$F1$ (1, 74) = 4.12, $p = 0.046$]. As for the between-subject factor, namely question wording, no effect emerged [$F1$ (1, 74) = 2.41, $p = 0.125$]. As for the first question, no interaction effects were found between any of the following factors: type of argument * wording [$F1$ (1, 74) < 1], connective * wording [$F1$ (1, 74) < 1], type of argument * connective [$F1$ (1, 74) < 1], type of argument * connective * wording [$F1$ (1, 74) = 1.34, $p = 0.251$].

4.5. Discussion

Before I discuss the results of the experiments presented in this paper in detail, I briefly recapitulate the two main findings. The results of the experiment targeting the perceived agreement between the speaker and the target of the straw man fallacies demonstrated a significant effect between statements containing a straw man and statements containing a non-fallacious reformulation. For both experimental versions (i.e., the between-subject factor, namely agreement and disagreement), participants attributed higher scores on the agreement-scale when there was no straw man present. The effect obtained on both questions. Overall, both questions on both versions of the experiment clearly demonstrated that the diverging opinions of the interlocutors in the dialogues were perceived more strongly in cases of fallacious arguments.

The results of the between-subject factor did not show any significant effects. There was no notable difference between the score relating to the type of argument (fallacious or non-fallacious) in the agreement and the disagreement version. Overall, this means that the question wording did not lead to significantly different results.

The results of the experiments presented in this paper were able to demonstrate that people are sensitive to the disagreement expressed through the use of a straw man fallacy. Throughout both experimental versions, participants were more likely to notice a disagreement when the speaker performed a straw man fallacy on their opponent. The fact that people actually notice the presence of a disagreement, and that they are aware of the fact that a straw man entails some kind of attack, supports previous findings. With the experiments in this paper, it has been shown that the people's preference for non-fallacious statements as shown in Schumann et al. (2019, 2021) also reflects their general capability of more easily detecting a disagreement when the speaker performs a straw man on his opponent. This result is due to the structure of the straw man itself, as the refutational dimension is expressed through the misrepresentational aspect (Aikin and Casey 2016; Oswald and Lewiński 2014). The fact that the attacker distorts the original content indirectly conveys their disagreement with the original position at the same time. In addition, a misrepresentation of the opponent's position puts an emphasis on the disagreement, as it not only shows that the interlocutors do not endorse the same position, but it also entails a critique towards the original position. Using such a fallacy has negative repercussions on the target of the straw man attack, as it conveys the impression that the target is less competent (see also de Saussure 2018; Oswald and Lewiński 2014). Through the refutation of the target's stance, the disagreement is further deepened, as it is implied that the position held by the target does not take into account an important aspect that is brought forward

by the misrepresented position of the attacker. Therefore, the results of the experiments indicate that the disagreement between the interlocutors might be perceived as deeper in cases of straw men compared to non-fallacious versions. It also emerged from the experiment that the overall agreement scores for the fallacious formulations were relatively high. In other words, people perceive that the person committing a straw man disagrees with their target, but they still attribute high scores on the agreement scale. This aspect might be related to the fact that the straw man fallacy appears to be effective in the sense of a winning argumentative move (see de Saussure 2018). The overall high agreement could therefore reflect the rhetorical competences of the speaker of the fallacy. However, in order to gain more insight into this aspect, it would be necessary to include a measure focusing on the reason why participants attribute such high agreement scores in cases of fallacious arguments. In sum, the experiments demonstrate that people are not only sensitive to manipulations concerning the misrepresentational dimension of the straw man fallacy as illustrated in previous work, but also to the refutational dimension of the straw man.

The experiments performed and presented in this paper were not able to replicate the effects of question wording in the context of perceived disagreement in straw man fallacies. The positive or negative wording of the experimental question did not lead to significant differences in perception. Different explanations for this result are possible. People's opinions might depend on the solidity of their mental context, which is based on three aspects: the standard of judgement, the frame of reference, and the attitude (Cantril 1941; Rugg and Cantril 1942). According to the explanation offered by Rugg and Cantril (1942), people with a solid standard of judgement and frame of reference, i.e., people with a strong set of values and rigid principles, are less likely to be influenced by the wording of questions when asked about their attitudes towards the topic at stake. In these cases, their own attitudes are more strongly based on higher values that are less permeable to external influences such as the positive or negative orientation of questions.

Another aspect that might explain the lack of effect found for the wording of the questions in the present experiment could be anchored in the design of the scales used to evaluate the statements. The evaluation was based on 6-point Likert scales, which ensured that participants had the possibility of giving a more nuanced opinion. Overall, this enabled more variability across the responses. The results obtained from the experiments presented in this paper can be linked to previous studies that considered the scale-size in respect to the question wording, such as Kamoen et al. (2011). From their results (see Section 3), it appears that a larger scale leads to more nuanced effects compared to binary scales. This would explain why the wording effect did not emerge in the experiment using 6-point Likert scales.

A third aspect that might explain the absence of an effect can be found in the experimental task. The participants were asked to evaluate the agreement or disagreement between two interlocutors. This means that it was not about the participants' own opinions and attitudes towards an expressed content, but that they were asked to judge the attitude of someone else. Typically, the studies conducted by Cantril (1940) or Kamoen (2012) targeted the participants' own attitudes towards the presented content. In the experiment presented in this paper, the participants were not personally involved in the opinion-building but had to evaluate the opinion built by the person performing a straw man and the attacked person. The provided explanations give us some direction as to how to interpret the lack of wording effect in the experimental context discussed in these experiments, but further investigations need to be conducted in order to more accurately explain the obtained results.

5. Conclusions

In conclusion, this paper has shown that people are aware of the refutational dimension of the straw man. The results clearly indicate that people notice the disagreement between the interlocutors when a speaker misrepresents their opponent's position and performs a straw man fallacy on them. The experiments presented in this paper have shown that people

perceive disagreement regardless of the way the experimental questions are formulated. Even when the formulation of the questions was oriented in a more positive way, focusing on the perceived agreement between the interlocutors in the dialogue rather than the disagreement, people were aware of the underlying attack and noticed the disagreement between both parties. This demonstrates that people notice the underlying function of the straw man, which is one of attack and refutation. This finding is reassuring and illustrates that people perceive the dichotomous characteristics encoded in a straw man fallacy. People are therefore not easily fooled and more perceptive than one might presume (see also Mercier 2020). On a more general level, the experiments conducted in this paper provide a more complete picture of the straw man. They show that, in addition to the people's ability to detect the presence of straw man fallacies and the importance of the linguistic formulation for their acceptability, there is an awareness linked to the refutational dimension as well. It thus demonstrates that both characteristics of the straw man can be tested in experimental settings. However, the experiments on the refutational dimension of the straw man also illustrate the need for further investigations into aspects such as the role of question wording for the refutational dimension of the straw man fallacy.

On a more general level, research as presented in this paper shows that approaches combining methodologies from experimental pragmatics and argumentation can be fruitful for a better understanding of the way fallacies such as the straw man are perceived. This type of study provides a more fine-grained analysis of specific factors such as the perception of disagreement or the wording effect on the basis of empirical data. It shows that a variety of aspects can be put to the test for empirical validation. Overall, this paper contributes to the research on straw man fallacies and shows the immense potential of experimental approaches to the study of fallacies in general.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to the non-invasive nature of the study and the resulting absence of risks.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the author.

Conflicts of Interest: The author declares no conflict of interest.

Notes

1 The experiment also targeted a third variable, namely the role of the causal connective *puisque* (closest to English *since*) as a linguistic marker highlighting the disagreement. In Schumann et al. (2019, 2021), using *puisque* to introduce a fallacious argument raised the awareness towards the fallacious content and consequently led to lower acceptability rates for straw men. In addtion, this connective conveys a tacitly negative attitude (Schumann et al. 2021; Franken 1996; Zufferey 2014). On the basis of these aspects, I argued that *puisque* could reinforce a disagreement expressed by straw men. The results of the experiments did not support this assumption. Since this variable was part of the experimental design, the materials and the means of all conditions including this dimension will be presented in order to provide a complete picture of the study, but it will not be discussed further in the analysis or the discussion.

2 This section reports the means and standard deviations for all conditions, but only the type of argument and the wording will be addressed further.

References

Aikin, Scott. 2019. Deep Disagreement, the Dark Enlightenment, and the Rhetoric of the Red Pill. *Journal of Applied Philosophy* 36: 420–35. [CrossRef]

Aikin, Scott, and John Casey. 2016. Straw Men, Iron Men, and Argumentative Virtue. *Topoi* 35: 431–40. [CrossRef]

Aikin, Scott, and John Casey. 2022. *Straw Man Arguments*. London: Bloomsbury Academic.

Aikin, Scott, and Robert Talisse. 2020. *Political Argument in a Polarized Age. Reason and Democratic Life*. London: Polity Press.

Bishop, George, Hans-Jürgen Hippier, Norbert Schwarz, and Fritz Strack. 1988. A Comparison of Response Effects in Self-Administration and Telephone Surveys. In *Telephone Survey Methodology*. Edited by Robert Groves, Paul Biemer, Lars Lyberg, James Massey, William Nicholls and Joseph Waksberg. Hoboken: Wiley, pp. 321–39.

Bizer, George, Shirel Kozak, and Leigh Ann Holterman. 2009. The persuasiveness of the straw man rhetorical technique. *Social Influence* 4: 216–30. [CrossRef]

Cantril, Hadley. 1940. Experiments in the wording of questions. *Public Opinion Quarterly* 4: 330–32. [CrossRef]

Cantril, Hadley. 1941. *The Psychology of Social Movements*. Hoboken: Wiley.

de Saussure, Louis. 2018. The Straw Man Fallacy as a Prestige-Gaining Device. In *Argumentation and Language—Linguistic, Cognitive and Discursive Explorations*. Edited by Steve Oswald, Thierry Herman and Jerôme Jacquin. Berlin: Springer International Publishing, vol. 32, pp. 171–90. [CrossRef]

Finocchiaro, Maurice. 2011. Deep disagreements: A meta-argumentation approach. In *Argumentation: Cognition and Community. Proceedings of the 9th Conference of the Ontario Society for the Study of Argumentation (OSSA), 8–11 May 2011, University of Windsor, Canada*. Edited by Frank Zenker. Windsor: OSSA Conference Archives, pp. 1–38.

Franken, Nathalie. 1996. Pour une nouvelle description de puisque. *Revue Romane* 31: 1–17.

Kalton, Graham, Martin Collins, and Lindsay Brook. 1978. Experiments in Wording Opinion Questions. *Applied Statistics* 27: 149. [CrossRef]

Kamoen, Naomi. 2012. *Positive versus Negative. A Cognitive Perspective on Wording Effects for Contrastive Questions in Attitude Surveys*. Utrecht: Universiteit Utrecht.

Kamoen, Naomi, Bregje Holleman, Pim Mak, Ted Sanders, and Huub van den Bergh. 2011. Agree or Disagree? Cognitive Processes in Answering Contrastive Survey Questions. *Discourse Processes* 48: 355–85. [CrossRef]

Lewiński, Marcin. 2011. Towards a Critique-Friendly Approach to the Straw Man Fallacy Evaluation. *Argumentation* 25: 469–97. [CrossRef]

Mercier, Hugo. 2020. *Not Born Yesterday. The Science of Who We Trust and What We Belive*. Princeton: The Princeton University Press.

Oswald, Steve, and Marcin Lewiński. 2014. Pragmatics, cognitive heuristics and the straw man fallacy. In *Rhetoric & Cognition: Theoretical Perspectives and Persuasive Strategies*. Edited by Thierry Herman and Steve Oswald. Bern: Peter Lang, pp. 313–43.

Rugg, Donald. 1941. Experiments in Wording Questions: II. *Public Opinion Quarterly* 5: 91–92. [CrossRef]

Rugg, Donald, and Hadley Cantril. 1942. The wording of questions in public opinion polls. *The Journal of Abnormal and Social Psychology* 37: 469–95. [CrossRef]

Schuman, Howard, and Stanley Presser. 1977. Question wording as an independent variable in survey analysis. *Sociological Methods & Research* 6: 151–70. [CrossRef]

Schumann, Jennifer. 2022. The Pragmatics of Straw Man Fallacies. An Experimental Approach. [Unpublished doctoral dissertation], University of Bern, Bern, Switzerland.

Schumann, Jennifer, Sandrine Zufferey, and Steve Oswald. 2019. What makes a straw man acceptable. Three experiments assessing linguistic factors. *Journal of Pragmatics* 141: 1–15. [CrossRef]

Schumann, Jennifer, Sandrine Zufferey, and Steve Oswald. 2021. The linguistic formulation of fallacies matters: The case of causal connectives. *Argumentation* 35: 361–88. [CrossRef]

Shields, Matthew. 2018. On the Pragmatics of Deep Disagreement. *Topoi* 40: 999–1015. [CrossRef]

Talisse, Robert, and Scott Aikin. 2006. Two Forms of the Straw Man. *Argumentation* 20: 345–52. [CrossRef]

Tseronis, Assimakis. 2021. How Face Threatening Are Disagreement Moves? A Proposal for an Integration of Insights from Politeness Theory into Argumentation Theory. In *The Language of Argumentation*. Edited by Ronny Boogaart, Henrike Jansen and Maarten van Leeuwen. Berlin: Springer International Publishing, vol. 36, pp. 167–87. [CrossRef]

van Eemeren, Frans, and Rob Grootendorst. 1992. *Argumentation, Communication, Fallacies. A Pragma-Dialectical Perspective*. Hillsdale: Lawrence Erlbaum Associates.

van Eemeren, Frans, and Rob Grootendorst. 2004. *A Systematic Theory of Argumentation: The Pragma-Dialectical Approach*. Cambridge: Cambridge University Press.

van Eemeren, Frans, and Peter Houtlosser. 1999. Strategic maneuvering in argumentative discourse. *Discourse Studies* 1: 479–97. [CrossRef]

van Eemeren, Frans, Rob Grootendorst, and Francisca Snoeck Henkemans. 2002. *Argumentation: Analysis, Evaluation, Presentation*. Mahwah: Lawrence Erlbaum Associates.

Walton, Douglas, Chris Reed, and Fabrizio Macagno. 2008. *Argumentation Schemes*. Cambridge: Cambridge University Press.

Zufferey, Sandrine. 2014. Givenness, procedural meaning and connectives. The case of French puisque. *Journal of Pragmatics* 62: 121–35. [CrossRef]