

Argumentation Quality Assessment for Persuasive Essays Annotation Guidelines

Santiago Marro[†], Serena Villata[†], and Elena Cabrio[†]

[†]Université Côte d’Azur, CNRS, Inria, I3S, France

1 Introduction

The objective of these annotation guidelines is to characterize different quality features of argument components applied to persuasive essays. More precisely, we use the dataset of persuasive essays annotated by Stab and Gurevych [1] with argument components together with the relations (i.e., support or attack) between the identified component in the essays.

Persuasive essays are meant to persuade the reader into supporting the stance the author is taking regarding a certain topic. Persuasive essays generally start with an introduction that describes the topic of the essay and rarely includes arguments, but the thesis statement expressing the stance of the author. We refer to this as Major Claim. The arguments supporting or attacking the major claim are contained in the paragraphs after the introduction. The last paragraph usually concludes the essay and contains a re-statement of the major claim.

An example of an essay can be found in Example 1. Figure 1 shows an example of an argument structure.

Taking advantage of the annotations done by Stab and Gurevych [1], we also propose to use the argument graph to assist annotators in their annotation process. Three quality attributes will be annotated, namely **cogency**, **reasonableness** and **argumentation rhetoric**.

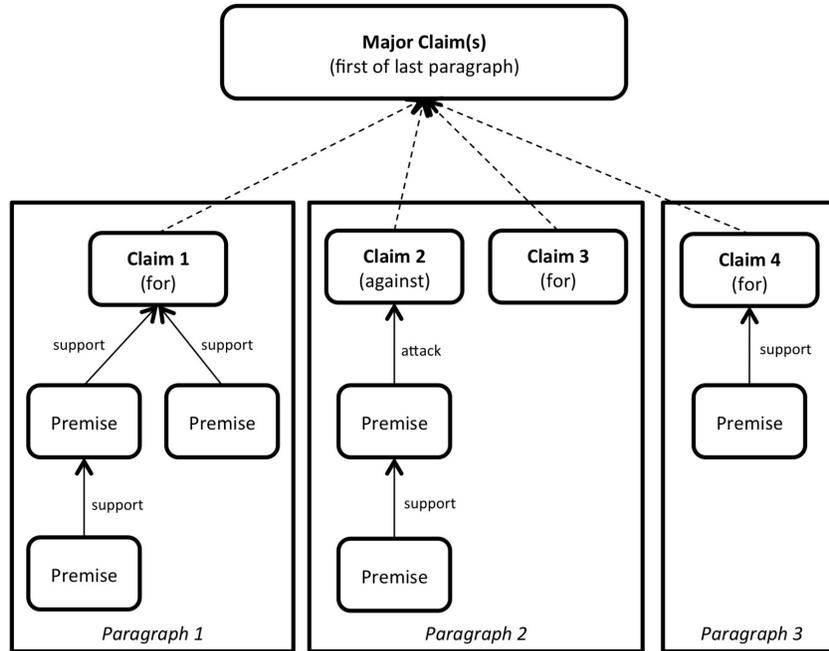


Figure 1: An example of the argumentation structure with its relations of a persuasive essay by Stab and Gurevych [1]

Persuasive Essay 1 *Should students be taught to compete or to cooperate?*

It is always said that competition can effectively promote the development of economy. In order to survive in the competition, companies continue to improve their products and service, and as a result, the whole society prospers. However, when we discuss the issue of competition or cooperation, what we are concerned about is not the whole society, but the development of an individual's whole life. From this point of view, I firmly believe that we should attach more importance to cooperation during primary education.

First of all, through cooperation, children can learn about interpersonal skills which are significant in the future life of all students. What we acquired from team work is not only how to achieve the same goal with others but more importantly, how to get along with others. During the process of cooperation, children can learn about how to listen to opinions of others, how to communicate with others, how to think comprehensively, and even how to compromise with other team members when conflicts occurred. All of these skills help them to get on well with other people and will benefit them for the whole life.

On the other hand, the significance of competition is that how to become more excellence to gain the victory. Hence it is always said that competition makes the society more effective. However, when we consider about the question that how to win the game, we always find that we need the cooperation. The greater

our goal is, the more competition we need. Take Olympic games which is a form of competition for instance, it is hard to imagine how an athlete could win the game without the training of his or her coach, and the help of other professional staffs such as the people who take care of his diet, and those who are in charge of the medical care. The winner is the athlete but the success belongs to the whole team. Therefore without the cooperation, there would be no victory of competition. Consequently, no matter from the view of individual development or the relationship between competition and cooperation we can receive the same conclusion that a more cooperative attitudes towards life is more profitable in one's success.

Steps for the annotation process

1. Read the argument component provided as well as its argument tree (the other argument components connected by argumentative relations to this component).
2. Establish if you can make an objective judgement; if not, please select the *cannot judge* label.
3. Assess the *Cogency* (this quality attribute is further detailed in Section 3):
 - Score it based on the following scale: 0, 10, 15, 20, 25.
4. Evaluate *Reasonableness*. This quality attribute will be further detailed in section 4:
 - You will score it based on the following scale: 0, 10, 15, 20, 25.
5. Evaluate the *Argumentation rhetoric*. This quality attribute will be further detailed in section 5:
 - Annotate it with one of the following labels “Logos”, “Pathos”, “Ethos”.
6. Continue with the next example.

2 Annotation for persuasive essays

In this annotation process, we will be annotating the quality of *argument components*. Arguments in their simplest form consist of two components, *claims* and *premises*. The claim represents the statement the author is trying to convince the reader to support to, and the premises can be seen as justifications for this claim. An argument could also be made with premises that aim at refuting a claim. In a typical argument structure, a thesis statement, also known as *major claim*, is presented followed by several *claims* that justify the author stance regarding the *major claim*. *Premises* then *support* or *attack* these claims providing the evidence necessary to persuade the reader. Argument components

and relations compose what we called the argument graph of an argument. An example of an argument graph is shown in Figure 1

Following the work of Wachsmuth et al. [2], different existing theories and approaches have been proposed in the literature to assess logical, rhetorical, and dialectical quality dimensions of arguments, deriving a systematic taxonomy. The three main characteristics Wachsmuth et al. [2] define within this taxonomy are the same we are annotating, namely **Cogency**, **Effectiveness** and **Reasonableness**, with the exception for Effectiveness, where we characterize this quality dimension with respect to the rhetoric aspects of the argument. For the rest of the guidelines we will refer to this as the **Argumentation Rhetoric**. In order to annotate Cogency and Reasonableness, Wachsmuth et al. [2] proposed 6 sub dimensions that we will not evaluate but their definitions will be integrated in our own definitions of said characteristics. Given that our objective is to annotate persuasive essays, we also want to take into account how these essays are assessed by teachers and professors. In the work by Stapleton et al. [3] the authors discuss and propose a rubric for persuasive writing that integrates the assessment of both argumentative structural elements and reasoning quality in students' persuasive writing by analyzing argumentative essays made by 125 students in Hong Kong. This rubric proposed by Stapleton et al. [3] contemplates several of the same characteristics present in our definition of Cogency and Reasonableness such as Relevancy and Acceptability as well as the presence of counterarguments and rebuttals.

For this we propose two different analytic scoring rubrics, one for Cogency and the other for Reasonableness. For Argumentation Rhetoric, we propose to annotate the rhetoric aspects of the argument in its whole following the rhetorical strategies defined by Aristotle, namely Logos, Ethos, and Pathos.

3 Cogency

We label Cogency taking into account the Analytic Scoring Rubric for Cogency in Table 1 based in the work by Stapleton et al. [3]. A scale of 0, 10, 15, 20, 25 is given for each way of characterizing the Cogency of a given argument.

The cogency dimension is focused on the *premises* of the argument to be labeled. As mentioned in Section 2, arguments in their simplest form are composed of claims and premises.

Following Table 1, we define the **acceptable** premises as the ones that are worthy of being believed, i.e., if you rationally think they are true or if you see no reason for not believing that they may be true.

The premises should be seen as **relevant** if they contribute to the acceptance or rejection of the argument's conclusion, i.e., if you think they are worthy of being considered as a reason, evidence, or similar regarding the conclusion.

Score: 25	Score: 20	Score: 15
a. Provides multiple reasons for the claim(s), and b. All reasons are sound/acceptable and free of irrelevancies	a. Provides multiple reasons for the claim(s), and b. Most reasons are sound/acceptable and free of irrelevancies, but one or two are weak	a. Provides one to two reasons for the claim(s), and b. Some reasons are sound/acceptable, but some are weak or irrelevant

Score: 10	Score: 0
a. Provides only one reason for the claim(s), or b. The reason provided is weak or irrelevant	a. No reasons are provided for the claim(s); or b. None of the reasons are relevant to/support the claim(s)

Table 1: Analytic Scoring Rubric for assessing Cogency.

We now describe the cogency dimension through some examples from the persuasive essays dataset [1] with the help of Figures 2 and 3. Starting with Figure 2 and following the Analytic Scoring Rubric for Cogency in Table 1 we annotated its cogency as Score: 25 (higher score). The author presents multiple premises (C, D and E) which are acceptable, relevant and sufficient to draw a conclusion. On the other hand, the example in Figure 3 was annotated as Score: 0 (lower score). The author provides no acceptable premises for their claim.

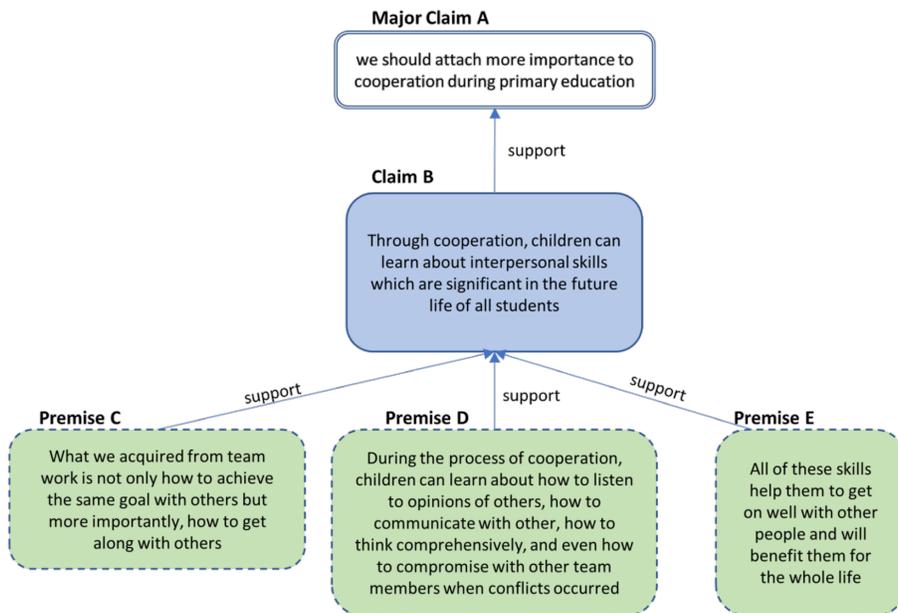


Figure 2: Argument graph example extracted from an essay consisting of one claim (in blue) and its premises (in green).

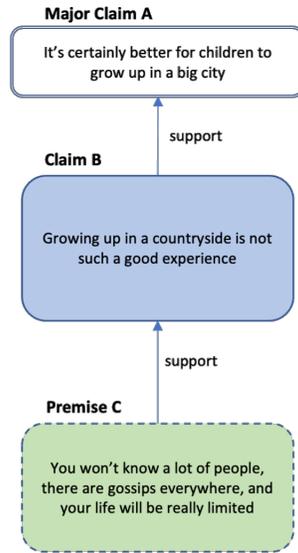


Figure 3: Argument graph example extracted from an essay from the persuasive essays dataset [1].

4 Reasonableness

Similar to what we did with Cogency in section 3, we follow again the Analytic Scoring Rubric for Reasonableness proposed by Stapleton et al. [3] in Table 2. This rubric has the argumentative components and their relations reported in the rows, and the quality of the supporting reasons in the columns. A scale of 0, 10, 15, 20, 25 is used for the two categories. Notice that Reasonableness will only be annotated if there are counterarguments, represented as Claim(s) attacking the Main Claim, or Claim(s) attacking other Claim(s), present in the persuasive essay. This evaluation is done in an argumentation graph level.

	Score: 25	Score: 20	Score: 15
Counterargument Data /Supporting Reasons for Alternative Point(s) of View	a. Provides multiple reasons for the counterargument claim(s) /alternative view(s), and b. All counterarguments/reasons for the alternative view(s) are sound/acceptable and free of irrelevancies	a. Provides multiple reasons for the counterargument claim(s)/alternative view(s), and b. Most counterarguments/reasons for the alternative view(s) are sound/acceptable and free of irrelevancies, but one or two are weak	a. Provides one to two reasons for the counterargument claim(s) /alternative view(s), and b. Some counterarguments/reasons for the alternative view(s) are sound/acceptable, but some are weak or irrelevant
Rebuttal Claim(s)	a. Refutes/points out the weaknesses of all the counterarguments, and b. All rebuttals are sound/acceptable c. The reasoning quality of all the rebuttals are stronger than that of the counterarguments	a. Refutes/points out the weaknesses of all the counterarguments, and b. Most rebuttals are sound/acceptable, but one or two are weak c. The reasoning quality of most rebuttals are stronger than that of the counterarguments, while one or two are equal to that of the counterarguments	a. Refutes/points out the weaknesses of all the counterarguments, and b. Some rebuttals are sound/acceptable, but some are weak c. The reasoning quality of some rebuttals are stronger than that of the counterarguments, while some are weaker than that of the counterarguments

Table 2: Analytic Scoring Rubric for Reasonableness

	Score: 10	Score: 0
Counterargument Data /Supporting Reasons for Alternative Point(s) of View	a. Provides only one reason for the counterargument claim(s)/alternative view(s), or b. The counterargument/reason for the alternative view is weak or irrelevant	a. No reasons are provided for the counterargument claim(s)/alternative view(s); or b. None of the reasons are relevant to/support the counterargument claim(s)/alternative view(s)
Rebuttal Claim(s)	a. Refutes/points out the weaknesses of some counterarguments, or b. Few of the rebuttals are sound/acceptable; most of them are weak, or c. The reasoning quality of most rebuttals are weaker than that of the counterarguments	a. No rebuttals are provided; or b. None of the rebuttals can refute the counterarguments

Figure 4 shows an example of an argument with counter-arguments and how we annotate it. Following the Analytic Scoring Rubric for Reasonableness in Table 2 we annotated the quality of the Counterargument Data as Score: 0 (lower score) and the quality of the Rebuttal Claim as Score: 25 (higher score). The author presents a single Counterargument (Claim E) with no premises supporting it. The rebuttal (Claim F) is sound/acceptable refuting the counterargument.

5 Argumentation Rhetoric

For this quality dimension, we will be annotating at the argument level which rhetoric strategy, defined by Aristotle, the argument is following: Logos, Ethos

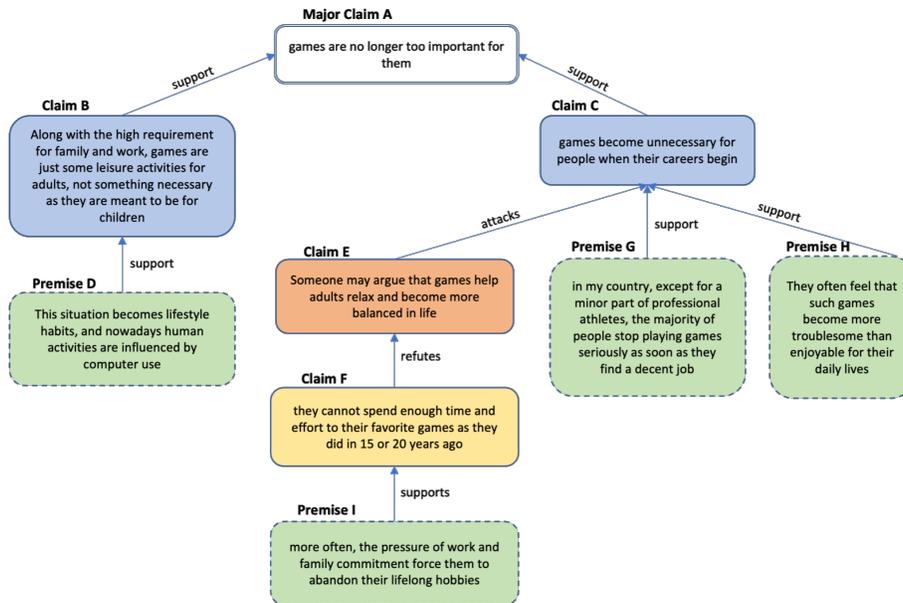


Figure 4: Argument graph example of an essay from the persuasive essays dataset [1]. Consists of two claims supporting the Major Claim (in blue), one claim (in red) attacking another Claim, one claim refuting a claim (in orange) and its premises (in green).

or Pathos.

Logos is the act of appealing to the audience through reasoning or logic. In Rhetoric (translated by Roberts, 2004), Aristotle defined logos as, “the proof, or apparent proof, provided by the words of the speech itself... by means of persuasive arguments” (p.7). To use logos would be to cite facts and statistics, historical and literal analogies, and citing certain authorities on a subject. The focus of logos should be well formed by a well-prepared writer so that it conducts a logical argument and evidence about it, which is applicable to a current situation.

Ethos or the ethical appeal is appealing to the audience through the credibility of the writer’s beliefs or authority. An author would use ethos to show to his audience that he is a credible source and is worth listening to. Ethos can be developed by choosing language that is appropriate for the audience and topic (also means choosing proper level of vocabulary), making yourself sound fair or unbiased, introducing your expertise or pedigree, and by using correct grammar and syntax.

Pathos or the emotional appeal, means to persuade an audience by appealing to their emotions. It could be defined as Aristotle explained: “putting the audience into a certain frame of mind. Persuasion may come through the hearers, when the speech stirs their emotion” (Roberts, 2004, p.7). Authors use

pathos to invoke sympathy from an audience; to make the audience feel what the author wants them to feel. A common use of pathos would be to draw pity from an audience. Pathos can be developed by using meaningful language, emotional tone, emotion-evoking examples, stories of emotional events, and implied meanings.

Taking into account these definitions, the goal of the annotation of the argumentation rhetoric dimension is to annotate an argument with one or more of the following labels: “Logos”, “Pathos”, “Ethos”. If no label can be assigned, the annotator will select the “no rhetorical strategy identified” label.

We will show now some examples about the annotation of the argumentation rhetoric in Figure 5 and Figure 4. In the first example, the author employs the Pathos structure. It appeals to emotions when she is describing in Claim B how “people are better taken care”. In Premise D “healthy workers and create more productivity” and how it helps to “save an amount of time as well as cost”, mentioned in Premise F.

For the second example, we can refer again to Figure 4, specifically on Claim C, where the author presents a credible argument with clear language and no emotional appeal is present. The author tries to persuade the reader by referring to her personal experience and knowledge, falling under the definition of Ethos. Therefore, we annotate this example as “Ethos”.

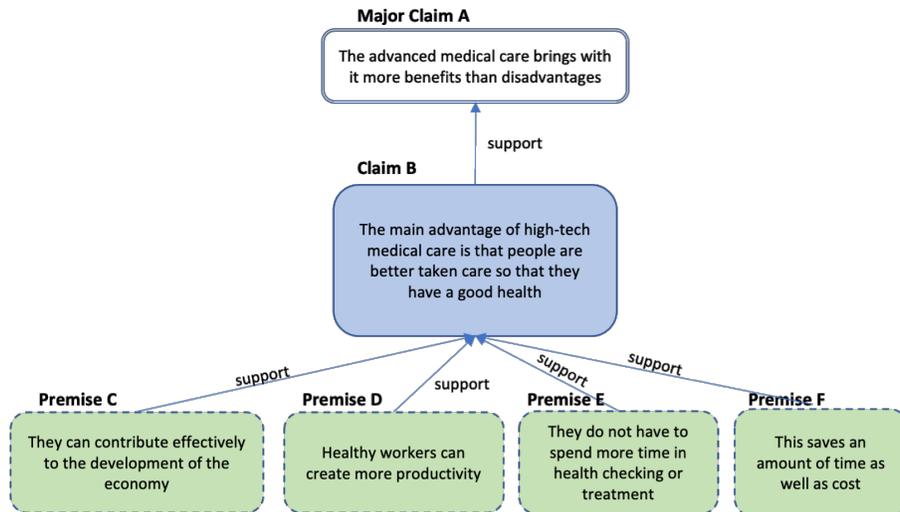


Figure 5: Argument graph example extracted from an essay.

References

- [1] Christian Stab and Iryna Gurevych. Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3):619–659, September 2017.
- [2] Henning Wachsmuth, Nona Naderi, Yufang Hou, Yonatan Bilu, Vinodkumar Prabhakaran, Tim Alberdingk Thijm, Graeme Hirst, and Benno Stein. Computational argumentation quality assessment in natural language. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 176–187, Valencia, Spain, April 2017. Association for Computational Linguistics.
- [3] Paul Stapleton and Yanming (Amy) Wu. Assessing the quality of arguments in students’ persuasive writing: A case study analyzing the relationship between surface structure and substance. *Journal of English for Academic Purposes*, 17:12–23, 2015.