The Importance and Use of Normative Criteria to Manipulate Argument Quality

Hans Hoeken, Jos Hornikx & Yvette Linders

To cite this article: Hans Hoeken, Jos Hornikx & Yvette Linders (2019): The Importance and Use of Normative Criteria to Manipulate Argument Quality, Journal of Advertising, DOI: 10.1080/00913367.2019.1663317

To link to this article: https://doi.org/10.1080/00913367.2019.1663317
The Importance and Use of Normative Criteria to Manipulate Argument Quality

Hans Hoeken  
_Utrecht University, Utrecht, the Netherlands_

Jos Hornikx  
_Radboud University Nijmegen, Nijmegen, the Netherlands_

Yvette Linders  
_Utrecht University, Utrecht, the Netherlands_

Argument quality plays an important theoretical and methodological role in persuasion research. Researchers frequently refrain from employing independent normative criteria to manipulate argument quality. Instead, they use pretests to qualify arguments that evoke predominantly favorable thoughts as strong and arguments that predominantly evoke unfavorable thoughts as weak. In this article, we analyze weak arguments as they have been used in actual studies. These weak arguments ranged from arguments referring to less favorable consequences compared to their strong counterparts, to consequences that are irrelevant to the participants, or even to undesirable consequences, thereby essentially functioning as counterarguments. We discuss the implications of this practice for our understanding of the persuasion process and our ability to provide evidence-based guidelines for message designers. We also provide guidelines on how to manipulate argument quality using normative criteria.

Argument quality is theorized to determine the outcome of the persuasion process if people are able and motivated to scrutinize the message (Petty and Cacioppo 1986). Argument quality is also manipulated by researchers to establish whether people have engaged in argument scrutiny. To that end, participants are exposed to either a message containing strong arguments or one containing weak arguments in support of the same claim, and then are asked to indicate their acceptance of the claim. If claim acceptance is found to be higher for the strong arguments message compared to the weak arguments message, participants must have been paying attention to the arguments. In the absence of such an effect, they have not. In a meta-analysis, Carpenter (2015) found 134 studies that manipulated argument quality and measured claim acceptance. The results show that strong arguments indeed yielded more positive attitudes than weak arguments, especially (but not exclusively) for participants who processed the message centrally. Combined with other reviews (Johnson et al. 2004; Park et al. 2007), these results appear to provide support for the importance of argument quality in the persuasion process.

In this research note, we take a closer look at how researchers have manipulated argument quality. Strong arguments typically refer to desirable attributes of the advertised product or desirable consequences of using it, examples being “Sparkling Garden eliminates 99% of germs on hands when used as a hand soap” (Jo 2004); “The VCR also...
includes a deluxe digital, on-screen timing program that determines how much tape is left, how much time is left in the current program, and how long the program has been playing” (Wheeler, Petty, and Bizer 2005); and “Comfrin lasts up to 3 hours longer than other aspirins” (Tormala, Briñol, and Petty 2006). The corresponding weak arguments in these studies show greater diversity: “Sparkling Garden eliminates 60% of germs on hands when used as a hand soap”; “The VCR also includes an eject button on its front face that permits you to remove the video and get a rough idea of how much tape is left”; and “Comfrin lasts almost as long as other aspirins.”

These manipulations of argument quality have been developed using an empirical approach to argument quality. In this approach, strong and weak arguments are identified by having participants rank arguments from strong to weak or report their thoughts when reflecting on the arguments. The weak arguments mentioned were found to elicit more negative thoughts than their strong counterparts and were thus all bona fide weak arguments from this perspective.

But one can also take a normative approach to argument quality. In this approach, argumentation theorists propose criteria that an argument should meet to be considered strong. From a normative perspective, the three previously mentioned weak arguments end up in different categories. The hand soap would be considered a weak argument as it brings about the desirable effect of eliminating germs to a lesser extent; the VCR argument would be considered irrelevant because all VCRs have an eject button, as a result of which it does not distinguish this VCR from its competitors; the painkiller argument would be considered a counterargument as it reveals that the product performs its main function more poorly than the competition.

We argue that normative considerations should be taken into account when manipulating argument quality for three goals: to better understand the role played by argument quality in the persuasion process, to prevent the development of unrealistic experimental materials, and to provide practitioners with evidence-based guidelines on what constitutes a strong argument. We also provide guidelines on how to manipulate argument quality using normative criteria. Although our analysis is intended to be quite general, for reasons of clarity we focus on one specific kind of argument: argument from consequences.

**EMPIRICAL AND NORMATIVE DEFINITIONS OF ARGUMENT QUALITY**

Argument quality is a central concept in the Elaboration Likelihood Model. Petty and Cacioppo (1986) employ an empirical method to identify strong and weak arguments. First, they developed a list of arguments that intuitively were considered strong or weak. Next, they had participants rate these arguments for persuasiveness. Finally, they had a different set of participants list their thoughts when reflecting on the highest and the lowest rated arguments. Strong arguments were defined as arguments evoking predominantly favorable thoughts when reflected upon, whereas weak arguments evoke mainly unfavorable thoughts. Favorable thoughts can be positive thoughts or negative thoughts depending on whether the claim propagates performing (“Buy this brand!”) or refraining from a behavior (“Stop smoking!”). A strong argument in favor of buying a certain brand evokes mainly positive thoughts about the behavior, whereas a strong argument against smoking evokes mainly negative ones. This approach is an empirical approach to argument quality relegating the answer to the question as to what is a strong or a weak argument to the perceptions of participants. It is also the most used method to manipulate argument quality; Carpenter (2015) reported it being employed in almost 90% of the studies in which argument quality had been manipulated. Petty and Cacioppo’s (1986, p. 32) definition of argument quality—strong arguments evoke predominantly favorable thoughts when reflected upon, whereas weak arguments elicit mainly unfavorable ones—is also often cited in research (e.g., see Batra and Stayman 1990, p. 203; Martin, Lang, and Wong 2003, p. 59).

O’Keefe and Jackson (1995) criticize this empirical approach. They state that it does nothing for our understanding of the role of argument quality in the persuasion process, as it remains silent on what aspects of strong arguments are responsible for the predominantly favorable thoughts or what characteristics yield the mainly unfavorable thoughts in case of weak arguments. To increase our understanding of what the active ingredients of argument quality are, O’Keefe and Jackson (1995) argue that “an independently-motivated account of argument quality” (p. 91) is required. Such an account would focus on standards or criteria for argument quality, thus introducing the normative approach to argument quality. In a normative approach, strong arguments are defined as meeting the normative criteria to a larger extent than weak arguments. The next section describes these normative criteria.

**NORMATIVE CRITERIA FOR ARGUMENT QUALITY**

O’Keefe and Jackson (1995) point to informal logic as a promising source of normative criteria for argument quality. Informal logic originated from the desire to improve people’s critical thinking skills, an important one being the ability to judge the value of arguments encountered in real-life policy debates, editorials, and ads (Blair
Informal logic thus aims to identify the norms that can be applied to assess the quality of everyday arguments. The triad acceptability, relevance, and sufficiency has been proposed as relevant norms for this purpose. Acceptability refers to accepting the content of the argument as true, for instance, that the hand soap would eliminate 99% of the germs. If this content is rejected, the argument will be considered weak. Relevance is about the adequacy of the link between the argument and the claim. The consumer may accept that the VCR has an eject button but, given that all VCRs have such a button, it cannot serve as an argument to claim that the consumer should choose this particular one. Sufficiency refers to the question whether the argument is sufficient to accept the claim. For instance, is the argument that a painkiller lasts up to three hours longer than other painkillers sufficient to accept the product’s superiority?

To assess differences on the third criterion, sufficiency, guidance is provided by the argument scheme approach (Blair 2015). An argument scheme is a template for certain recurring argument types, such as argument from example (in which a claim is supported by reference to a specific case) and argument from authority (referring to the opinion of an expert). For each argument type, criteria have been proposed to assess the strength of such arguments. Arguments that satisfy these criteria to a stronger extent are considered strong(er). Different argument types have different associated criteria. For example, an argument from authority needs to satisfy criteria (e.g., expertise, trustworthiness) that are different from those applicable to an argument from example (e.g., number, typicality).

In domains such as advertising, health communication, and political communication, the single most common type of argument is the argument from consequences, both in communication practice (Schellens and de Jong 2004) and in academic research (O’Keefe 2013b). Walton (1996) defines arguments from consequences as “a species of practical reasoning where a contemplated policy or course of action is positively supported by citing the good consequences of it. In the negative form, a contemplated action is rejected on the grounds that it will have bad consequences” (p. 75). In the argument from consequences, the claim that a certain action should be carried out or a specific product bought is supported by arguments referring to desirable consequences of that action. The argument from consequences can also refer to attractive product attributes (e.g., the brand being cheaper, a laptop having more memory capacity, or a food product being produced in a more environmentally friendly way) that will lead to desirable consequences (e.g., less money spent, more information can be stored, or a better environment) when choosing the product. For instance, in the strong arguments mentioned in the introduction, these consequences refer to an almost complete disinfecting of one’s hands, knowing how much video tape is left for recording, and being pain free for a longer period of time.

There is consensus within argumentation theory (e.g., see Schellens and de Jong 2004; Walton 1996) and persuasion theory (e.g., see Petty and Wegener 1991) on the two most important normative criteria that an argument from consequences has to meet to be considered strong. First, the consequence referred to has to be desirable (in case of an argument in favor of the behavior) or undesirable (in case of an argument against the behavior). Second, the consequence has to be likely to result from performing the behavior. The more desirable and the more likely the consequence is, the stronger the argument in favor of the behavior is; the more undesirable and the more likely the consequence is, the stronger the argument against a certain action is.

**ACTUAL ARGUMENT QUALITY MANIPULATIONS AND NORMS**

From a normative perspective, a strong argument from consequences meets the desirability and probability criteria to a larger extent than a weak argument. Inspection of experimental materials reveals that the strong arguments typically refer to desirable consequences or attributes (such as products being cheaper or more effective, or taxes being spent on important issues). In some cases, researchers have used weak arguments that indeed refer to a (slightly) less desirable attribute compared to the strong arguments. For instance, Raju, Unnava, and Montgomery (2009, Study 3) developed a weak arguments ad for a car stating “3-year full-service warranty, 4 out of 5 stars on independent crash test results, and 25 miles per gallon in the city”; in the strong arguments ad, the car fares better on each of these attributes: “10-year full-service warranty, 5 out of 5 stars on independent crash test results, and 35 miles per gallon in the city.” In Jo’s (2004) strong argument, the hand soap eliminates 99% of germs; in the weak argument it eliminates only 60% of germs.

However—and this is the main point of this research note—some of the so-called weak arguments are not weak arguments from a normative perspective. Instead, the putatively weak arguments turn out to be either irrelevant arguments or counterarguments. Arguments are irrelevant if they refer to attributes or consequences that play no part in the target audience’s decision process or do not distinguish the product from its competitors; arguments are counterarguments if they refer to attributes or consequences on which the advertised product does worse than its competitors or if they depict the advertised product’s performance as unsatisfactory with respect to some important attribute or consequence.
In some studies, researchers have used weak arguments that refer to consequences or attributes that do not meet the criterion of relevance. For instance, Clark and Thiem (2015) use the following argument to sell a detergent: “The packaging of phosphate detergents is specially designed to be more visually appealing ... you might leave them out in the open as pieces of art.” Calanchini, Moons, and Mackie (2016, Study 4) argue that a tax increase is needed to beautify highways as they are currently visually boring. Helweg-Larsen and Howell (2000) referred to a condom’s fashionable colors in their ad, and Escalas (2004) employed the following weak argument to sell a running shoe: “Westerly running shoes have never been tested on laboratory animals.” In these cases, the arguments fail to be relevant either because the desirable consequence does not factor into the evaluation process (e.g., detergents are bought for their cleaning power, not for the aesthetic appeal of their packaging) or because it does not distinguish it from its competitors (e.g., no running shoes are tested on animals).2

In other studies, researchers have used weak arguments that, upon closer examination, turn out to be counterarguments. For example, in some weak argument messages, arguments are included in which the advertised product does not function well on an essential characteristic (e.g., Sanbonmatsu and Kardes’s (1988) ad for a pen read: “Writs legibly with only an occasional skip”) or in which the advertised product is negatively compared to its competitors on an important attribute (e.g., Priester and Petty’s (2003) ad aimed to sell Rollerblades by stating that they were “15% to 20% more expensive than other roller-blades”). A rather striking example is the weak arguments message developed by Chu and Kamal (2008): “The most important reasons to choose Yottabyte Laptop are the use of the obsolete technology, the mediocre quality of components, the heaviest materials, the dimmest screen, and the average lasting battery.”33

Arguments such as the ones about the detergent’s packaging or mediocre quality of the laptop components are of course likely to evoke mainly unfavorable thoughts, which qualifies them as weak arguments when using the empirical definition of argument quality. From a normative perspective, however, these are not weak arguments but irrelevant or even counterarguments. In the next section, we address the undesirable consequences of using empirically based argument quality manipulations.

WHY USING NORMATIVE CRITERIA IS IMPORTANT

When manipulating argument quality, researchers have compared strong arguments to counterarguments, irrelevant information, or normatively weak arguments. These different comparisons have implications for (1) the realism of the resulting ad, (2) our insight into the role of argument quality in the persuasion process, and (3) our ability to provide evidence-based advice for practitioners on how to select strong arguments.

Irrelevant Information and Counterarguments Yield Unrealistic Messages

Using an empirical approach to argument quality, Petty et al. (2004, p. 79) argue that it does not matter on what normative criteria the strong and weak arguments differ from each other if the manipulation is used for the purpose of assessing which route participants followed; as long as the strong arguments evoke more favorable thoughts than the weak arguments when reflected on, the manipulation serves its purpose. To achieve that end, Vargas, Duff, and Faber (2017) argue that argument quality manipulations should not be too subtle.

However, including irrelevant information or counterarguments can result in unrealistic messages. Weaver, Hock, and Garcia (2016) argue it is unlikely that a professional “would think that including a specious or false argument would have persuasive appeal” (p. 30). It indeed seems unlikely that any advertising professional would try to sell a VCR by referring to its eject button or a laptop by lauding its obsolete technology. As a result, a difference in impact between messages with strong arguments and those including irrelevant information or counterarguments could be the result of the latter being perceived as an unrealistic, unprofessional effort. Petty and Cacioppo (1986, p. 32) themselves argue that the weak arguments should differ from the strong arguments only in strength but not in believability.

Using unrealistic weak arguments messages can complicate the interpretation of the effects of other variables. Argument quality manipulations are typically used in 2 × 2 experiments in which researchers are interested in what the impact is of the second independent variable and use argument quality to assess whether participants have been processing centrally or peripherally. Chu and Kamal (2008), for instance, manipulated not only argument quality but also the source’s trustworthiness. They found that for the strong arguments messages, the more trustworthy source is more persuasive than the less trustworthy source. However, for the weak arguments message, the opposite effect is reported: The less trustworthy source is more persuasive compared to the more trustworthy one. Why would a more trustworthy source be less convincing than a less trustworthy one? An explanation for this surprising effect may be that the “weak arguments” message containing counterarguments was considered as especially unrealistic coming from a trustworthy source. Why would a Massachusetts Institute of Technology alumnus who has won the Turing
award state that a laptop should be bought because of its obsolete technology and its dimmest screen? The unreality of the message makes it difficult to interpret the effect of this second independent variable.

Using Normative Criteria Enables Theoretical Progress

The ELM dominates the conceptualization of the persuasion process. In this model, central processing is described as seeing the merits of strong arguments and the flaws in weak ones (Petty and Cacioppo 1986, p. 182), as “effortfully assessing the merit of the arguments” (Petty et al. 2004, p. 70), and “considerable thought about weak or specious arguments” (Petty et al. 2005, p. 90; emphasis added). Carpenter’s (2015) meta-analysis on the impact of argument quality seems to provide evidence for this depiction of central route processing, as it concluded that the greater persuasiveness of strong arguments over weak arguments was more pronounced for participants who processed the message centrally.

However, Carpenter distinguishes between studies in which argument quality was approached empirically and those using a more principled approach. He concludes that “when researchers used thought listing to construct strong and weak arguments, the weak arguments seem to have been particularly likely to be rejected. Next, when message qualities were altered to produce strong and weak arguments rather than pretesting them, the argument quality effects tended to be smaller” (Carpenter 2015, p. 516).

This difference in effect size between these two approaches may be the result of the empirical approach leading to including irrelevant information and/or counterarguments in the weak arguments messages. Comparing messages containing such arguments to messages containing strong ones does not constitute a test of the ELM claim that during central processing people evaluate the merits of the arguments. To assess this central tenet of the model would require independent normative criteria. We are not the first to raise this issue (e.g., Petty et al. 2004; Mongeau and Stiff 1993; Jackson 1995). However, the finding that so-called weak arguments messages contain irrelevant information or even counterarguments, combined with Carpenter’s results that argument quality effects are notably stronger for empirically developed manipulations, strongly testifies to the importance of using normative criteria for our understanding of the persuasion process.

Using Normative Criteria for Evidence-Based Guidelines

Manipulating argument quality using normative criteria and assessing their impact is important for practical purposes as well. O’Keefe (2003) argues that “without effect-independent characterizations of message variations … message designers will have little guidance about the construction of effective messages” (p. 269). Translated to argument quality: For researchers to provide message designers with evidence-based guidelines for what arguments are more convincing than others and what standards for a specific type of argument are most important to meet, argument quality has to be manipulated using independent criteria.

Given the ubiquity of the argument from consequences in persuasive messages, it would be essential for message designers to know what makes an argument from consequences strong and convincing. There is ample evidence for the importance of the consequence’s desirability for the strength of an argument: the more desirable the expected consequence, the more convincing the argument is (see Hoeken, Timmers, and Schellens 2012; O’Keefe 2013b). The importance of the consequence’s likelihood is unclear. Whereas it is logical to assume that an unlikely consequence should have less impact than a more likely one, there is little empirical evidence for this effect (Hoeken et al. 2012; Van Enschot - Van Dijk, Hustinx, and Hoeken 2003; Hustinx, Van Enschot, and Hoeken 2007; O’Keefe, 2013b). For message designers, it is important to know whether, and/or under what conditions, they should invest in providing support for the likelihood of a consequence to occur.

HOW TO MANIPULATE ARGUMENT QUALITY?

How should researchers proceed when they want to manipulate argument quality? In a typical advertisement, product attributes (e.g., guarantee, price) and the consequences of product use (e.g., beautiful hair from a shampoo, better sound from a speaker) serve as such arguments. First, researchers must identify what the relevant arguments are for the specific product or behavior. This can be established either by analyzing a sample of advertisements for this type of product or by asking a sample of consumers to list the attributes and consequences that are relevant to them when choosing such a product. This strategy should preclude that a VCR’s eject button or a detergent’s package design are chosen as arguments.

Second, the strong and weak arguments should preferably refer to the same attribute (e.g., guarantee period) or consequence (e.g., percentage of germs eliminated by a hand soap) but differ in the extent to which they possess this attribute or deliver the consequence. A pretest will be needed to identify what people consider the average score for this attribute (e.g., 90% of germs eliminated, three-year guarantee) and what they consider to be the upper limit (e.g., 99% of germs eliminated, five-year guarantee, which could serve as a strong argument) and the lower limit (e.g., 60% of germs eliminated, one-year guarantee, which could serve as a weak argument).
This strategy is applicable to quantifiable attributes and consequences but not to categorical ones such as “beautiful hair” or “great sound.” In those cases, a source manipulation could be helpful. Advertisers have a vested interest in having people accept their arguments. If, however, an independent source, such as a consumer organization, stated that the shampoo leads to more beautiful hair, or that the TV displays the colors brilliantly, the argument becomes stronger. In a similar way may the (high or low) percentage of satisfied customers provide a strong or weak argument in favor of a product. In the latter case, a pretest will be needed to assess what participants would consider the average percentage of satisfied customers as well as the lower and upper limits of these percentages.

Third, researchers should measure the believability of the different ads in their studies. Especially in the case of the weak arguments messages it is important to check whether the manipulation yielded a message that is considered unprofessional, unrealistic, or unbelievable (see Geuens and De Pelsmacker 2017). If there are differences in believability between the message versions, mediation analyses may help assess whether and to what extent these differences played a role in the persuasive impact of these messages, enabling a more nuanced interpretation of the results.

CONCLUSION

Weak arguments messages in experiments on persuasion host a variety of arguments that from a normative perspective range from truly weak arguments in favor of the claim, via irrelevant information, to arguments against the message’s claim. Irrelevant information and counterarguments do not run the risk of being too subtle for centrally processing participants to notice and thus serve the methodological goal of detecting whether participants paid attention to the argumentative content. Yet including them may decrease the message’s believability, which results in a confound that makes it difficult to interpret the impact of other factors being manipulated in the study.

Employing independent normative standards to manipulate argument quality could prevent the development of unrealistic messages. In addition, using independent normative standards contributes to solving what Fishbein and Ajzen (1981) already noted as “probably the most serious problem in communication and persuasion research” (p. 351); our lack of knowledge about what constitutes a strong argument. More than 30 years later, O’Keefe (2013a) argues that identifying “message properties that enhance persuasiveness under conditions of high elaboration ... would represent an important advance in the understanding of persuasion generally and argument quality specifically” (p. 144). Employing independent normative criteria for argument quality could lead to this advance.

ACKNOWLEDGMENTS

The authors thank Hans Beentjes, Peter Jan Schellens, Daniel O’Keefe as well as the associate editor and two reviewers for their insightful comments on previous versions of this manuscript.

NOTES

1. If the audience is not convinced about the consequence’s desirability or the likelihood, this part of the argument becomes a contested claim. The advertiser may include arguments to support the claim that the consequence is desirable (or likely to occur as a result of this action). The arguments can be of various types (e.g., argument from authority, argument from example).

2. If there were running shoes that had been tested on laboratory animals, the argument that Westerly shoes does not do so could be relevant to those consumers who do not want to buy products that have been tested on laboratory animals. However, as long as all brands refrain from using laboratory animals, this argument is not relevant because it does not distinguish Westerly shoes from its competitors.

3. As one reviewer pointed out, the inclusion of counterarguments in a message might evoke the concept of two-sided messages. However, we think that the kinds of counterarguments used in weak arguments messages are not the same as the sorts of counterarguments one characteristically sees in research on one-sided and two-sided messages. A two-sided message is defined by Kamins and Assael (1987) as “a message that presents positive claims on important attributes, but downgrades or limits product or brand performance claims on attributes of minor significance to the consumer so as to establish credibility without deterring purchase” (p. 29). The weak arguments messages are not two-sided messages for two reasons. First, weak arguments messages typically lack positive claims about important attributes. The Yottabyte Laptop ad states—explicitly—that the laptop fares worse on all attributes mentioned. Similarly, the weak Comfrin painkiller ad also lacks any positive claims on its attributes. Second, Kamins and Assael (1987) state that the downgrade is “on attributes of minor significance to the consumer” (emphasis added). In the weak argument messages, the negative information is not on product attributes that are of minor significance. For instance, the major benefit of a painkiller is its effectiveness as well as the duration of this effect. Stating that this painkiller’s effect lasts for shorter periods than those of its competitors is thus not about a less important attribute; nor are statements that the laptop has the dimmest screen or that a pen skips. Apart from the benefits, there are the costs. Stating that a product is more expensive than its competitors is, again, negative information on an important attribute. None of the weak arguments messages that included information on the product being more expensive contained information on the superiority of the product on important benefits.
REFERENCES


