CrossMark

Laypeople's Evaluation of Arguments: Are Criteria for Argument Quality Scheme-Specific?

Peter Jan Schellens[1] · Ester Šorm[2] ·
Rian Timmers[3] · Hans Hoeken[4]

Abstract Can argumentation schemes play a part in the critical processing of argumentation by lay people? In a qualitative study, participants were invited to come up with strong and weak arguments for a given claim and were subsequently interviewed for why they thought the strong argument was stronger than the weak one. Next, they were presented with a list of arguments and asked to rank these arguments from strongest to weakest, upon which they were asked to motivate their judgments in an interview. In order to assess whether lay people apply argument scheme specific criteria when performing these tasks, five different argumentation schemes were included in this study: argumentation from authority, from example, from analogy, from cause to effect, and from consequences. Laypeople's use of criteria for argument quality was inferred from interview protocols. The results revealed that participants combined general criteria from informal logic (such as

✉ Peter Jan Schellens
p.j.schellens@let.ru.nl

Ester Šorm
e.sorm@tudelft.nl

Rian Timmers
r.timmers@has.nl

Hans Hoeken
j.a.l.hoeken@uu.nl

[1] Centre for Language Studies, Radboud University Nijmegen, P.O. Box 9103,
6500 HD Nijmegen, The Netherlands

[2] Centre for Languages and Academic Skills, Delft University of Technology, P.O. Box 5015,
2600 GA Delft, The Netherlands

[3] HAS Den Bosch University of Applied Science, P.O. Box 90108, 5200 MA 's Hertogenbosch,
The Netherlands

[4] Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, 3512 JK Utrecht,
The Netherlands

relevance and acceptability) and scheme-specific criteria (such as expertise for argumentation from authority, similarity for argumentation from analogy, effectiveness for argumentation from consequences). The results supported the conventional validity of the pragma-dialectical *argument scheme rule* in a strong sense and provided a more fine-grained view of central processing in the Elaboration Likelihood Model.

Keywords Argumentation schemes · Argumentation from authority · Argumentation from example · Argumentation from analogy · Argumentation from cause to effect · Argumentation from consequences · Evaluation criteria · Conventional validity · Central processing

1 Introduction

To assess the quality of an argument, criteria can be applied that are specific for the argument scheme at hand.[1] For instance, when evaluating an argument from analogy, different criteria are relevant compared to evaluating an argument from example or an argument from cause to effect. In this paper, we will focus on the question whether people can and do apply such argument scheme specific criteria. This question is relevant for the research within the social-psychological tradition of argumentation and persuasion as well as that within argumentation theory.

Within social-psychological research, two ways of processing persuasive messages are distinguished: peripheral versus central processing (Petty and Cacioppo 1986), or—in a slightly different way—heuristic versus systematic processing (Chaiken 1987; Chen and Chaiken 1999). According to Petty and Cacioppo's (1986) Elaboration Likelihood Model (ELM) receivers who lack the motivation and/or capacity to scrutinize the arguments will evaluate the acceptability of the claim employing less energy consuming procedures, for instance by applying heuristics such as 'if an expert says it, it is probably correct'. If they are motivated and able to evaluate a persuasive message thoroughly, they will critically evaluate the arguments comparing them to their own beliefs about the world as well as the extent to which these arguments should lead to adapting one's opinion.

However, it is not very clear what central processing of a persuasive message, or elaborating on a message, exactly entails. See, for instance, the following description by Petty and Priester (1994, pp. 98–99):

> The first, or "central route", to persuasion involves effortful cognitive activity whereby the person draws upon prior experience and knowledge in order to carefully scrutinize all information relevant to determining the central merits of the position advocated. (…) the message recipient under the central route is actively generating favorable and/or unfavorable thoughts in response to the

persuasive communication. The goal of this cognitive effort is to determine if the position advocated by the source has any merit.

Until now, the numerous experimental studies into the Elaboration Likelihood Model have not been able to provide a more detailed picture of what this effortful cognitive activity or careful scrutiny of arguments implies. If we can answer the question whether lay people apply scheme-specific criteria when thoroughly evaluating arguments, this will clarify what the concept of central processing can entail (Schellens and De Jong 2004).

From an argumentation theoretical point of view, the question of what central processing entails, is equally relevant. The issue of whether lay people would use argument scheme specific criteria when evaluating arguments speaks to the conventional validity of norms or rules developed in argumentation theory. In their normative theory, Van Eemeren and Grootendorst (1992, 2004) specify a system of rules for a critical discussion. To solve a dispute in a reasonable way, discussants have to follow these rules. For instance, the *rule of freedom* says that 'discussants may not prevent each other from advancing standpoints or from calling standpoints into question' (Van Eemeren et al. 2009, p. 21) and the *Relevance rule* runs as follows: 'Standpoints may not be defended by non-argumentation or argumentation that is not relevant to the standpoint' (Van Eemeren et al., p. 22).

For our study, the *Argument scheme rule* is relevant: 'Standpoints may not be regarded as conclusively defended by argumentation that is not presented as based on formally conclusive reasoning if the defense does not take place by means of appropriate argument schemes that are applied correctly' (Van Eemeren et al. 2009, p. 23). In pragma-dialectical theory, three categories of argumentation schemes are distinguished: argumentation based on a symptomatic relation, on a relation of analogy, and on a causal relation (Van Eemeren et al. 2002, pp. 96–102). More specific subcategories are distinguished within each category. For instance, argumentation from example is based on a symptomatic relation and argumentation from consequences is a subcategory of causal argumentation (Garssen 1997). Each argumentation scheme is accompanied by its own evaluation criteria or so called 'critical questions', sometimes further specified for subcategories. So, in a reasonable discussion the *Argument scheme rule* demands the use of appropriate argument schemes in a way that meets the requirements of the critical questions. A contribution or discussion move that violates one of the rules or requirements is (by definition) a fallacy.

A normative theory, like pragma-dialectical argumentation theory, has to meet at least two requirements: problem validity and conventional validity. Problem validity demands that the rules and norms for critical discussion 'are instrumental in fostering the resolving of a difference of opinion' (Van Eemeren et al. 2009, p. 25). In this paper we will not deal with this kind of validity but focus on the requirement of conventional validity. This requirement stipulates that 'the rules should also be acceptable to "the discussants" or to "ordinary language users"' (p. 27) without a special training in the analysis or evaluation of argumentation. Van Eemeren et al. indicate that 'it is an analytical-theoretical question whether the rules for critical

discussion are problem valid, but it is a matter of empirical investigation to determine whether the rules have conventional validity' (p. 27).

The requirement of conventional validity of pragma-dialectical rules was object of an extensive research project by Van Eemeren et al. (2009). In over twenty experiments, Van Eemeren et al. had their participants read short conversations, consisting of three utterances, in which the conversation partners had a discussion. For each conversation, two or more versions were construed which differed with respect to whether the discussion contribution violated a discussion rule (turning it into a fallacious discussion move) or followed this specific rule (making it a reasonable discussion move). Results of these experiments were highly consistent: almost always participants judged the fallacious moves as less reasonable than their reasonable counterparts, leading Van Eemeren et al. to conclude that the pragmadialectical rules are conventionally valid.

For the purposes of this paper, the experiments of Van Eemeren et al. (2009) related to the *Argument scheme rule* are especially relevant. They studied four fallacies in this field: the *argumentum ad consequentiam*, the *argumentum ad populum*, the fallacy of the slippery slope, and the fallacy of false analogy (pp. 163–199). The first and second type of fallacy are considered inappropriate argument schemes,[2] whereas the third and the fourth are incorrect applications of the argument schemes from consequences (or pragmatic argumentation) and from analogy. For these studies, the results are highly consistent as well: the fallacies are judged unreasonable and less reasonable than the non-fallacious applications of argumentation from consequences and from analogy.

At first, Van Eemeren et al. (2009) conclude that these results provide evidence for the conventional validity of the pragma-dialectical *argument scheme rule*. The results indicate that the *argument scheme rule* successfully predicted the judgments of laymen on the reasonableness of discussion moves. In the final chapter of their book, Van Eemeren et al. (2009, p. 220) pose the question what their results imply about laypeople's criteria: 'What sorts of norms lie behind their judgments?' And 'to what extent are these norms (…) in agreement with the rules in the pragma-dialectical ideal model?' In order to address these questions, Van Eemeren et al. asked participants to motivate their reasonableness judgments in a number of experiments. However, interpretation of these results proved very difficult. Systematically, the respondents (aged 15–17) 'found it an extremely difficult abstract question to answer why a discussion contribution is reasonable or not', resulting in 'massive non-response' (Van Eemeren et al. 2009, pp. 220–221). The study's design was unable to elicit the evaluation criteria used by the participants.[3] Van Eemeren et al. (2009, p. 222) conclude that 'the qualitative oriented study into the motivations of the judgment of reasonableness provides no clear-cut conclusions for the (degree) of conventional validity of the pragma-dialectical discussion rules'.

---

[2] In the case of the argument *ad populum*, this argument could also be regarded as an incorrect use of the argumentation from authority (Van Eemeren et al. 2009, pp. 182–183).

[3] The disappointing results of the motivation task in the study of Van Eemeren et al. (2009) could be the result of the participants' age (they were all high school pupils) and their task: they had to write down their motivations.

As can be gathered from the quote above, the conventional validity of a rule can be established to a stronger or weaker degree. In a weak form, the results reveal that the judgments of lay people correspond to the judgments reached if the rule is applied correctly but without evidence that lay people have actually used this rule. Conventional validity is established to a stronger degree if the lay people's judgments do not only correspond to the predicted judgments but also are the result of applying the rule or system of rules. The experimental studies of Van Eemeren et al. (2009) provide strong evidence for the weaker conventional validity of the argument scheme rule but do not warrant a stronger interpretation.[4]

The underlying reasoning of the participants in the Van Eemeren et al. (2009) studies remains unclear. One cannot rule out the possibility that participants applied more general rules to reach their reasonableness verdict, such as a general relevance criterion: the respondents could consider slippery slope arguments and false analogies less relevant and, thus less reasonable. Participants could also have some knowledge of schemes for frequent fallacies. As soon as they recognized such a fallacy in a discussion move, they could have lowered their reasonableness judgment. Although the results of Van Eemeren et al. (2009) provide evidence for the weak conventional validity of argument scheme specific criteria, the results do not enable a strong conventional validity conclusion. To increase our insight into this issue, research employing a different research design is needed.[5]

Social-psychological theories do not answer the question as to what central or systematic processing of persuasive messages exactly entails. Research on the conventional validity of fallacious and non-fallacious discussion moves has yet to uncover what kind of norms laymen use to evaluate arguments. In this study, we employ a research method that stimulates participants to centrally process arguments and to reflect upon the criteria they employ when doing so.[6] Such a method enables addressing the following research question:

> Are the criteria used by laymen to evaluate argumentation specific for the argumentation scheme involved?

Argumentation schemes may be seen as conventional ways of defending a position. Goddon and Walton (2007, p. 267) characterize argument schemes as 'stereotypical

---

[4] Van Eemeren et al. (2009) also relate the extent to which a norm can be considered as conventionally valid to the difference in effect size for the reasonableness scores of the fallacies compared to their non-fallacious counterparts: 'In relative sense, the larger the effect sizes associated with the fallacies covered by a pragma-dialectical rule, the more the claim to conventional validity is substantiated' (Van Eemeren et al. 2009, p. 222). In our view, this only makes sense in concern with the conventional validity in a weak sense. The effect sizes did not say anything about the underlying norms or criteria and so did not support the conventional validity of the pragma-dialectical rules in a strong sense.

[5] In a critical response-test, Garssen (1997) investigated whether critical comments of lay people (high school pupils) on argumentation can be related to critical questions for argumentation based on a symptomatic relation, a causal relation or a relation of analogy. That appeared to be the case for a relatively small part of the comments: 31.7% of the comments on argumentation from analogy was scheme-specific, as were 11.0% of the comments on causal argumentation and only 5.9% of the comments on argumentation from a symptomatic relation (Garssen 1997, p. 210).

[6] Based on the results of the critical response-test of Garssen (1997), we decided not to use a written task, not to recruit secondary school pupils as participants, and to use argument schemes more concrete than those distinguished in pragma-dialectics.

patterns of defeasible reasoning', Blair (2001, p. 375) regards argument schemes as 'familiar patterns of reasoning or argument', and Van Eemeren and Grootendorst (1992, p. 96) define them as 'more or less conventionalized ways of presenting the relation between what is stated in the argument and what is stated in the standpoint'. All these definitions suggest that argument schemes are conventionally used in discussions in ordinary language. Moreover, the results of Van Eemeren et al. (2009) suggest that ordinary language users do not only distinguish different patterns of argument but may also apply specific criteria for evaluating their strength.

Of course, we do not assume that knowledge of schemes and criteria has exactly the format of argumentation schemes and criteria formulated by argumentation scholars: a composition of partly formalised propositions, of which one represents a standpoint and one or more others represent the argument(s), accompanied by criteria in the form of critical questions. We hypothetically assume that the vocabulary of laymen in discussing the quality of arguments will reveal evaluation criteria that are not applicable to argumentation in general, but only to a limited subcategory of arguments.

Unfortunately, there is no generally agreed upon classification of argumentation schemes. Walton et al. (2008) list over sixty argumentation schemes (including fallacious schemes and some schemes for formal reasoning). It seems unrealistic to assume that lay people possess such an elaborate classification. In pragma-dialectical theory, only three, highly abstract argument schemes are distinguished. If we restrict our study to these three main categories, the evaluation norms are likely to be abstract as well. Argumentation based on symptomatic relations, for instance, consists of—among other subtypes—argumentation from authority as well as argumentation from example. In our view, it is likely that language users will not recognize these two argument types as one and the same and could use different criteria when evaluating them. Therefore, we will use the typology of Schellens (1985, also Kienpointner 1992, pp. 201–215) and have selected five uncontroversial argument schemes that are distinguished in most typologies, albeit sometimes with a slightly different label: argumentation from authority, from analogy, from example, from cause to effect, and from consequences.[7]

2 Method

To uncover what criteria lay people use for evaluating the quality of different types of argumentation, we opted for a primarily qualitative design. In an individual interview, participants were presented with two cases. In the first case, the so-called open case, they had to produce a strong and a weak argument of a particular type in support of a claim. Afterwards, they were asked to provide their reasons for considering the strong argument as stronger than the weak one.

---

[7] Selection of these argumentation schemes has also been based on their function as backing of the premises in pragmatic argumentation: as arguments to support the desirability or probability of consequences. For a more extended account, see Šorm (2010) and Timmers (2014).

Subsequently, they were presented with a second case: the closed case. Here, they received a number of arguments, all of the same type and in support of the same claim, and they were asked to rank these arguments from strongest to weakest. In the subsequent interview, participants were asked to explain what reasons they had for their ordering. The participants were not asked to what extent they accepted the claim; they only had to judge the quality of the argumentation.

We conducted a study to assess whether individual interviews or group interviews (so called focus groups, see Krueger and Casey 2000) yielded better results. This study showed that individual interviews were most productive and more practical (see Šorm et al. 2007).

2.1 Materials

Participants were told that they had applied for a job as a speech writer for a Minister. As part of the procedure, they had to show that they were able to distinguish strong from weak arguments. To that end, they had to come up with a strong and a weak argument of a certain type. They were also told beforehand that after providing the arguments, they would have to provide their reasons why they considered the strong argument to be stronger than the weak argument. To increase the likelihood that respondents would come up with the intended type of argument, participants were asked to finish a specific sentence. For the argument from analogy, participants were asked to finish the sentence: "Because that is comparable to…" for both the strong and the weak argument. In the closed case, participants were presented with a claim and five to seven arguments of the same type. Participants were asked to rank these arguments from strongest to weakest. The Appendix contains the materials of the open and closed cases for the argument from analogy.

For each type of argument (argumentation from authority, from example, from analogy, from cause to effect, and from consequences) two open and two closed cases were constructed. This amounts to a total of ten open and ten closed cases.

2.2 Participants

A total of 108 participants took part. In the study on argumentation from authority, 24 students (18 women, 6 men) of a Dutch university participated. Their average age was 21, ranging from 18 to 28. For the other four argument types, visitors of the public library in a large city in the eastern part of the Netherlands acted as participants. For the studies on argumentation from analogy and argumentation from consequences, 22 participants took part. For the studies on argumentation from examples and from cause-to-effect, 20 participants took part in each of the studies. Their average age was 38, ranging from 15 to 70. The group consisted of 44 male and 40 female participants. Their level of education ranged from secondary school only to a master's degree. The participants received 10 euro for their participation. None of the participants had ever received formal training in argumentation or argumentation theory.

2.3 Procedure

The interviews concerning the argumentation from authority were conducted by master students of a communication program who had been trained in conducting interviews. In the other studies, one of the researchers approached the participants at the entrance of the library. They were asked if they were willing to participate in a study on governmental communication which would take approximately 30 min. Participants were randomly assigned to an open and a closed case of a specific type of argument. The interviews were subsequently conducted in a separate room in the library.

2.4 Data Analysis

Each interview was transcribed. The interview protocols were analyzed in three steps. First, the protocols were divided into parts. Each part was about the participants' consideration of why an argument was considered strong (or weak), or why it was considered stronger (or weaker) than another one. Care was taken that the parts included enough contextual information to understand the participant's utterances. Irrelevant parts (e.g., "This was the first case. Now we go on to the second one." Or "Can I read it again?") were left out.

Second, one or more criteria were ascribed to a single part. The wording of a criterion was kept close to the words used by the participants while at the same formulating them slightly broader so that it could be applied to other arguments of this type. The second part resulted in a list of criteria as they had been applied to that specific type of argument. In the third step of the analysis, different formulations of what appeared to be the same criterion were grouped together and a more encompassing label was suggested for that criterion.

For the second part of the analysis, an independent rater was involved. A total of 115 parts, equally distributed over the five types of argument, were labelled by an independent rater. The interrater reliability for the two raters was substantial (Cohen's κ ranged from .69 for the argument from cause to effect to .80 for the argument from authority).

3 Results

The results are presented separately for each argumentation scheme. Tables 1, 2, 3, 4 and 5 contain the criteria revealed in the open and closed cases and the number of participants that used a criterion. Subsequently, examples of criterion usage from the interview protocols will be provided.

3.1 Argumentation from Authority

Table 1 provides an overview of criteria for argumentation by authority identified in the interview protocols that have been used by at least five participants.[8] Plusses and minuses indicate that a criterion was used/not used in the open and/or closed cases.

---

[8] Other criteria (used by fewer than five participants): certainty of statement (3); external consistency (2); flexibility (2); ahead-of-time expert (1). For a more detailed presentation of results, see Šorm (2010) and

Table 1 Criteria for argumentation from authority

Criterion	Used in open cases	Used in closed cases	Number of participants using a criterion (n = 24)
Expertise*	+	+	24
Trustworthiness	+	+	22
Recency	−	+	13
Carefulness**	+	+	16
Internal consistency	+	+	5

* *Expert*, *Relevant expertise* and *Ability to provide evidence* combined

** *Hedged statement* in Šorm (2010)

Firstly, almost all participants used expertise and trustworthiness criteria to evaluate argumentation from authority. Among others, the following argumentation from authority was presented to the participants: 'Argument C: The well-known soccer players Ruud van Nistelrooy, Edgar Davids and Edwin van der Sar hold the opinion that that nuclear energy is the best way of energy production'.[9] The participant in example (1) does not accept this argumentation because she questions the expertise in nuclear matters of these well-known Dutchmen (P = participant; I = interviewer.)

(1) P: C, with support of well-known football players an argument is provided, while they, yeah, maybe they are knowledgeable on these issues, but they are not really specialized. So, I would not refer to them so easily. I do not think that that is really convincing.
I: Why do you think that it's not really convincing?
P: Because … they are football players. That's what they're known for. If they were also specialized in this field, then you have to add that to it: also specialized in nuclear energy.

In example (2), a participant positively evaluates the trustworthiness of a university professor in nuclear physics.

Footnote 8 continued

Timmers (2014). Because we do not aim at an experimental comparison of open and closed cases, we also abandon frequencies at this point. In general, the closed cases revealed more criteria than the open cases.

[9] Soccer players pleading in favor of nuclear energy could be considered as so far-fetched that participants would reject it without any other criterion in mind than an 'absurdity check' of relevance. However, prominent sports men and women are employed to sell all kinds of products in commercials; using them as quasi-experts in another field is more convenient than the example suggests. More important, in response to these kind of arguments, participants appear to apply scheme-specific criteria to qualify the argument as weak. One could take this as evidence for the accessibility of such criteria when evaluating even these 'absurd' cases.

(2)   I: Why do you consider that one the very best?
       P: Because he is independent; he has no vested interest in being in favor or
       against it. And because he is not associated with any organization that
       benefits from the use of nuclear energy, so … And the more independent, the
       better. So he can evaluate the advantages and disadvantages better. It is less
       subjective.

Secondly, a substantial number of participants applied the 'recency' criterion. For
instance, when evaluating a quote from Oppenheimer about nuclear power, dated in
the early fifties, they raised the question whether that opinion was still up to date.

(3)   P: Fine that he developed that atom bomb, but meanwhile … We're not in the
       early 50 s anymore, but quite some time later. And I don't know whether you
       can be bothered to stay up to date on the issue.

Thirdly, several participants applied criteria when confronted with a source using
hedges to qualify his or her opinion (e.g., 'Under certain circumstances nuclear
energy could help'). Such a nuanced opinion led to more favorable opinions for some
of the participants:

(4)   P: Yes, when I read him, he made the most trustworthy impression on me, the
       most convincing as well because, let me think, if you look at A, it's stated
       very simply: they think it can be reversed and that's it. Then there it is still
       slightly more specific, that under certain conditions, and with the help of
       nuclear… that is somewhat more conditionally put. That simply sounds more
       reliable than simply stating: that will work with…

Finally, the criterion of internal consistency was used by a minority of participants.
These participants rejected an argument in favor of nuclear energy in which the
authority referred to was the chair of Greenpeace in the Netherlands. These
participants considered the opinion expressed as inconsistent with previous
statements on this issue.

3.2 Argumentation from Analogy

Table 2 shows the criteria used by five or more participants to evaluate
argumentation from analogy.[10] The vast majority of participants applied criteria
specific to this type of argumentation: they evaluated the overall similarity of the
cases compared or—more specifically—the similarities and/or dissimilarities of
these cases. The questions whether there are sufficient relevant similarities or
whether the dissimilarities are negligible, can be regarded as subordinated to the
criterion of overall similarity. In example (5) a participant formulates this criterion
as follows:

---

[10] Other criteria (used by fewer than five participants): comprehensibility (4); permissibility (3);
concreteness (3); objectivity (1); support (1).

Table 2 Criteria for argumentation from analogy

Criterion	Used in open cases	Used in closed cases	Number of participants using a criterion (n = 21)
Overall similarity	+	+	18
Relevant similarities	+	+	20
Irrelevant dissimilarities	+	+	19
Acceptability analogon	+	+	17
Relevance	+	+	12
Acceptability	+	+	8

(5) P: Yes, especially how the comparison is built and how the two, the two elements in the comparison how they agree. That's what I paid most attention to. (…) Yes, whether they are alike, whether those two things can be compared. Whether they are in essence the same, so to say, or that you're not comparing apples to oranges.

Applying the second criterion, participants were looking for similarities of the analogue cases; applying the third criterion, they were looking for dissimilarities. When such (dis)similarities were identified, participants assessed their relevance to the issue at hand.

A majority of participants tried to verify the probability or acceptability of the crucial property in the compared case.

(6) P: Because, ehm, it has to do with the current situation in which we live, actually. It has been decided that many harmful products, which if they find out the damage they do, that it is just that they get taxed, usually because they are harmful too.

The participant explains his own strong argument (from the open case): in comparable cases taxes have been introduced by referring to the harmful consequences of the taxed products; since calorie rich snacks also have harmful consequences, a tax is justified.

In (7), an argumentation from analogy is criticized that says that sustainable energy is cheap because of governmental financial support, therefore the sustainable building of rental houses should be financially supported as well.

(7) P: Yes, actually because I tripped over the claim that sustainable energy is cheap. That seems to me as categorically untrue. So that's why you put that one in the final position.

In this case, the participant does not base his (negative) evaluation on the comparison but rather on the fact he considers a statement to be untrue. Cases such

as these are considered to reflect the application of a general, not scheme-specific criterion: acceptability.

In a similar fashion several participants applied a general criterion of relevance[11]:

(8) P: Because it, I think, does not address what the issue is. He could get AOW, but then for a reduced amount. And it is more an argument for a severe reduction of his AOW, but not to abolish it.

3.3 Argumentation from Example

Table 3 contains the criteria used by at least five participants to evaluate argumentation by example.[12]

A preliminary observation is that for this argumentation scheme only few criteria were identified in the open cases. The participants rarely came up with a strong and a weak argumentation by example. The claims which they had to support, were of the form 'If measure M is implemented, then effect E will occur' (for instance: 'If employers will contribute more to day care, employees will more easily combine their jobs and care for their children'). Although the task was to give (weak and strong) examples in support of this claim, the participants came up primarily with other advantages and disadvantages, respectively as strong and weak arguments for measure M. The most frequently used criteria concern the relevance and causality of the argumentation from example.[13]

In the following fragments, participants evaluate argumentation in favor of the claim 'If you choose a Sciences study, you will make good money in the future'. The arguments from example were amongst others: 'Peter Philips studied veterinary medicine and he has a nice job' and 'Janneke Oorthuys did a Psychology degree and earns an above-average income'.

(9) P: (…) Peter Philips. Yes, he has a very nice job and, ehm, (…) It's got nothing to do with it. He has a nice job, okay, but that says nothing about how much he earns.

---

[11] In studies into informal logic, both argumentation schemes accompanied with scheme-specific critical questions and general evaluation criteria for reasonable argumentation have been formulated. Acceptability, relevance and sufficiency are the most frequently mentioned (see, e.g., Govier 2005, pp. 63–64 and Johnson and Blair 2006). In this view relevance is a gradual quality: an argument can be more or less relevant. In pragma-dialectical theory relevance is a binary concept: the *rule of relevance* requires that an argument has to be relevant. Only if that criterion is met, the argument qualifies for evaluation on the base of the *argument scheme rule*.

[12] Other criteria (used by fewer than five participants): typicality (3); action desirability (3).

[13] Initially, in the design of this study all argumentation schemes were related to pragmatic argumentation: the open and closed cases of argumentation by example supported a cause-to-effect statement about a measure and his consequences. Of course, this restricts the generalizability of the criteria found, especially of the criterion 'causality of example'.

Table 3 Criteria for argumentation from example

Criterion	Used in open cases	Used in closed cases	Number of participants using a criterion in closed cases (n = 20)
Relevance example	–	+	17
Causality example	–	+	17
Number of examples	+	+	9
Acceptability example	–	+	9
Concreteness	–	+	7

(10) P: (…) on the final position, I had the, ehm, the one who studied psychology, because yeah, that's just not a Sciences program.
I: HmmHmm
P: And yeah, that's just got nothing to do with it, so to say.
I: What does it have nothing to do with?
P: With ehm that the sciences lead to a high salary.
I: HmmHmm
P: Because it is simply not about a sciences study, so yeah, then ehm, it's just not an argument.

In both fragments, participants apply the relevance criterion: in fragment (9), the argument is considered weak because the effect mentioned ("nice job") is considered irrelevant to the claim ("making good money"); in fragment (10), the relevance of the cause is questioned as Psychology is not considered to be a Sciences program. The frequent evaluation of the relevance of examples seems to replace the theoretical criterion of representativeness or typicality of examples. However, it is difficult to decide if the criterion in these cases is a scheme-specific requirement to argumentation from example or a general, not scheme-specific, criterion of relevance, applicable to all kinds of argumentation.

The frequent application of a criterion of causality of the example most probably was an artefact of the material used in this study: it does not seem a useful criterion for all argumentation from example. It is useful only where causal generalizations are supported by examples, as was the case in our material. For instance, this is very clear in the evaluation of the next argumentation from example, 'Philip Freriks, the anchorman of the NOS daily news show did a master in chemistry and he is in the top 10 of richest Dutch people':

(11) P: (…) it seems to me that there is not really a relation between studying chemistry and ending up at the daily news, so ehm. I mean would rather a Journalism program, or something like that, yeah, it don't think that you need to do chemistry for that.

The participant, considering that following a chemistry program is not the most probable cause of becoming an anchorman, critizes the causality of the example.

Less frequently the number and the acceptability of examples were evaluated. If an argument consisted of three examples, it was often evaluated more positively than an argument containing one example.

(12) P: Argument J I like better, well actually also because there are more, ehmm, (…) because all others contain only one.
I: Okay, and why do you think that's stronger, that there are more?
P: Ehm … yeah, because then there are more examples of, ehm, people for whom that argument holds, so to say. Ehm, well, the more people the argument applies to, ehm yeah, the more likely it is that it holds for you too.

The acceptability of the example was evaluated rather frequently. In such cases a participant indicated that an example was (probably) not true.

(13) P: Philip Freriks, well I, I think he's a languages guy. And I do not believe he did Chemistry.

3.4 Argumentation from Cause to Effect

Table 4 shows the criteria used by at least five participants to evaluate argumentation from cause to effect.[14] In this part of our study only a minority of participants (six out of twenty) succeeded to formulate weak and strong arguments of the relevant type. As a consequence, only a few criteria came to light in the open cases. Often a weak argument formulated to support the claim 'If measure M is taken then (desirable) effect E will occur' resulted in an argument referring to an undesirable consequence of taking measure M, thereby resulting in an argument against taking measure M rather than a weak argument in support of it. Similar patterns surfaced in the closed cases.

In the closed cases almost all participants applied the sufficiency criterion, asking themselves whether the cause mentioned in the argument was sufficient to evoke the effect predicted. For instance, the argument: 'Investing in schools for youth theatre generally leads to a sustainable future', was presented as an argument in support of the claim 'Giving more money to schools for youth theatre will promote the quality of society'. The participant in example (14) expressed doubts about the sufficiency of the cause:

(14) P: Yeah, I found that complete nonsense. (…) It will probably have a small contribution of course, but I do not think that, that it will really lead to a sustainable future. It will contribute a little bit, but, yeah, I didn't consider him very strong. It's not like society will immediately and completely change because more money is put in arts and culture, schools for youth theatre, I think.

---

[14] Other criteria (used by fewer than five participants): cause necessity (4); effect desirability (3).

Table 4 Criteria for argumentation from cause to effect

Criterion	Used in open cases	Used in closed cases	Number of participants using a criterion in closed cases (n = 20)
Sufficiency of cause	+	+	19
Relevance of cause or argument	−	+	7
Sufficiency and necessity of cause	−	+	6

Other criteria appeared to be used less frequently to evaluate argumentation from cause to effect. Relevance of the argument or a part of the argument was used by a minority of participants. For instance, this criterion is at stake when the participant does not see any relation between the cause in the argument and the cause mentioned in the conclusion:

(15)   P: And argument A, ehm, there I think the relation between the two claims, reasons, is completely lost. "Employees who have been ill for a long time and get time off, will in general return faster to their normal job." That says nothing at all. (…) "So", like it's a conclusion (…). Those two sentences have nothing in common. That's why Argument A is in final position".

In this and similar cases, the participants possibly applied a general criterion of relevance that can be used regardless of the argumentation scheme actually used.

Finally, several participants appeared to apply a necessary cause criterion in evaluating argumentation from cause to effect. Apparently, they read the argument as a plea for implementing measure M, rather than as a prediction that M will have effect E. As a result, they reasoned what (other) implications Measure M might have and depending on the (un)desirability of these consequences, evaluated the desirability of Measure M.

3.5 Argumentation from a Desirable Consequence

Argumentation from consequences or pragmatic argumentation appears in many forms. One of the elementary forms is to claim the desirability of a measure or action by referring to the desirable effect this action or measure will have. Table 5 contains the criteria used by five or more participants.[15]

The criteria 'desirability of effect' and 'seriousness of problem' were used by all participants. These criteria are closely related. In example (16) the participant evaluated the argument that a social work placement for adolescents will help solve

---

[15] Other criteria (used by fewer than five participants): comprehensibility (4); completeness (3); concreteness (2); counterarguments (1).

Table 5 Criteria for argumentation from advantageous consequence

Criterion	Used in open cases	Used in closed cases	Number of participants using a criterion (n = 22)
Desirability of effect	+	+	22
Seriousness of problem	+	+	22
Effectivity*	+	+	20
Efficiency	+	+	7
Costs/disadvantages	+	+	6
Relevance	+	+	7
Acceptability	+	+	5

* *Effectivity* and *Support of effectivity* summed

the problems in certain sectors suffering from severe personnel shortage. Although the argument is presented in a problem-solving frame, the participant formulates his criticisms in terms of the desirability of the effect.

(16)   P: Yeah, yeah, I find that more relevant to the interest of rather, not so much to that of the adolescents themselves. (…) Filling shortages on the job market, that does not seem to me, yes, it is nice that it happens, but it does not seem to me in the direct interests of the adolescents.

In example (17), a participant expresses doubts about the seriousness of the problem.

(17)   P: It says that reduction of emissions will result in a slower warming of the earth, and then the conclusion is drawn that a reduction of the emission of harmful substances is desirable, whereas it does not say why, what's wrong with global warming. (…) That has to be included for it to be a good argument. (…). Well alright, on the other hand people know that global warming is a problem. So, from that perspective it is not the worst argument either.

At an abstract level, both examples concern the evaluation of the effect. The effect is sometimes framed as an advantage of an action and sometimes as contributing to the solution of a problem.

Almost all participants applied the effectiveness criterion to evaluate a proposed action. In the next two examples the participants doubted if a social work placement will reduce the number of drop outs from school.

(18)   P: Because I am wondering whether it has much impact on finishing school or not, that work placement. I do believe that it helps, but whether that is the decisive factor determining whether someone will drop out or not, that, I can't imagine, that it will have such an impact.

Two other scheme-specific criteria, efficiency (Is this the most efficient way to reach effect E?) and costs or disadvantages (Is this measure not too expensive or come

with serious disadvantages?) were less frequently used. In example (19), a participant evaluates the argument that social work placements will help reduce personnel shortages. He mentions two disadvantages.

(19)   P: But I also think, actually, that it is unfair competition for people working there or who could work there. I don't think it's such a good argument. Ehm, yeah, I believe that if you, for instance, force someone of 16 years old to choose from a limited number of sectors, for instance, taking care of senior citizens or so, well, and how motivated these people will be who have to do that. I don't think that's always good, I guess.

Finally, also argumentation from consequences was evaluated by applying not scheme-specific criteria of relevance and acceptability. The participant in example (20) evaluates the acceptability of an argument that says that private clinics will create more efficiency in health care.

(20)   P: because it is a generally accepted fact. Everyone will agree with it, or at least, most people. Ehm, yeah, that's the reason I considered it strong.

4 Conclusion and Discussion

Are the criteria used by lay people to evaluate arguments specific for the argumentation scheme involved? The answer is at least partly affirmative. The results reveal that lay people use scheme-specific criteria for the evaluation of each of the argumentation schemes studied. Participants evaluated argumentation from authority employing criteria of expertise and trustworthiness of the source; argumentation from analogy was evaluated taking into account the similarities and differences between the cases compared; in the evaluation of argumentation from examples, the number of examples was considered; the quality of argumentation from cause to effect was considered to be related to cause sufficiency; argumentation from consequences was judged employing the consequences' desirability and the efficiency of means. All of these criteria correspond with criteria that have been suggested in the argumentation literature (see for instance Freeley and Steinberg 2000, Garssen 1997, Kahane 1992, Meany and Shuster 2002, Warnick and Inch 1989, Reinard 1991, Schellens 1985, Walton 1996).[16]

Although not as often as scheme-specific criteria, general informal-logic criteria were employed as well. Criteria as acceptability and relevance, as proposed in informal logic (Govier 2005, Johnson and Blair 2006) were used in the evaluations of four of the five argument schemes. General criteria were used frequently by participants evaluating cases of argumentation from analogy and consequences as well as sometimes when evaluating argumentation from examples and from cause to effect. General informal-logic criteria were not used for the evaluation of argumentation from authority; for that type of argumentation, only scheme-specific criteria were used.

---

[16] Šorm (2010) and Timmers (2014) present an overview of evaluation questions from argumentation literature for each of the five argumentation schemes studied.

The results obtained could be influenced by the materials employed, the participants' interpretation of the task, and the method of analysis of the interview protocols. In the closed case, participants received several arguments of the same kind and were forced to compare them. This design could have led participants to formulate criteria that they had never used before. However, this seems unlikely. For instance, if a participant had never considered the relevance of a source's expertise when evaluating an argument from authority, he or she would have had to (1) identify the difference in expertise when comparing two arguments differing on this aspect, (2) realize that this difference was relevant to the argument's quality, (3) compute the implications for the quality of the two arguments at hand, and (4) express these qualifications during the interview. The immediacy and fluency with which participants expressed their opinions and evaluations during the interviews suggest that no such complex reasoning preceded their reactions.

A second question is how the participants interpreted the task of ranking arguments from strong to weak. One interpretation is to consider a strong argument meeting normative criteria to a larger extent than a weak argument. A different interpretation would be to consider the argument that would be most likely to convince the audience as the strongest one, regardless of the extent to which this argument meets normative criteria. In a preliminary study focusing on the best design for obtaining this type of data, participants appeared to have interpreted the instruction as distinguishing between arguments that meet normative criteria to a larger or a lesser extent (Šorm et al. 2007). The results from the interviews in the study at hand reveal that the participants, again, focused on normative criteria to distinguish stronger from weaker arguments. Participants appear to have interpreted the task as focusing on the normative quality of arguments.

A third methodological issue is to what extent the interpretation of the interview protocols was biased towards the identification of scheme-specific norms. We cannot rule out such a bias. Hardly ever, a participant formulated a criterion explicitly; therefore the analysts had to derive a criterion from the participant's comments on the individual arguments. For instance, in response to a (weak) argument from authority, one participant supported his judgment by stating "he doesn't have a clue" whereas another said: "he doesn't have the education for this". Both comments were classified as referring to the *expertise* criterion. Similarly, in response to an argument from cause to effect, one participant said "That doesn't strike me as very likely" whereas another one responded to the same argument as "It seems to me that there is more needed for that (to happen)". Both statements were categorized as an application of *sufficiency of cause*—even though none of the participants used this label. Had participants been using a simpler, collective standard (such as, "This is a stupid argument"), their responses would not have been classified as reflecting the *expertise criterion* or the *cause sufficiency criterion*. The fact that frequently general, non-specific, criteria were identified appears to indicate that the raters held an open eye for criteria other than scheme-specific ones. Nevertheless, we refrain from formulating conclusions about the relative frequency of general and scheme-specific criteria. The main conclusion from our qualitative study is that both scheme-specific and general criteria were used.

What are the implications of our results for the conventional validity of the *argument scheme rule*? The results of Van Eemeren et al. (2009) already provided support for the conventional validity of this rule to the extent that lay people's judgments about an argument's reasonableness were in line with those reached by application of the normative criteria implicated by the rule. However, it was unclear whether the lay people's judgments were the result of applying such criteria. Our results show that lay people indeed use scheme-specific criteria next to general informal criteria.

We do not hypothesize that all language users have the same argumentation schemes and evaluation criteria at their disposal. Undoubtedly, knowledge of schemes and criteria will vary with cognitive capacity, education, profession, and experience. Still, the number of participants applying the same criterion can serve as an indication of the conventional validity of that criterion. The following criteria have been applied by at least 85% of the participants: the criteria *expertise* and *trustworthiness* (for argumentation from authority), *similarity* (for argumentation from analogy[17]), *example relevance* (for argumentation from example[18]), *sufficiency of cause* (for argumentation from cause to effect) and *desirability*, *problem seriousness* and *effectivity* (for argumentation from consequence). The widespread use of these criteria by highly heterogeneous sets of participants is a strong indicator of the conventional validity of these criteria. If a criterion is used by fewer participants, this could imply that it is conventionally less valid in a given community or it is valid in a smaller community. It is also possible, however, that participants know the criterion but did not feel the need to apply that criterion because they had already used another criterion to support their ranking of strong and weak arguments. The interviews were not about evoking an exhaustive set of criteria so it is possible that certain infrequently mentioned criteria are still widely known by lay people but that they did not have to use them to carry out the task at hand.

Apart from being relevant to the conventional validity of the argument scheme rule, the results also speak to the issue of what central processing in het Elaboration Likelihood Model could entail. Petty and Priester (1994, pp. 98–99) define central processing as 'determining the central merits of the position advocated' through a careful evaluation of the arguments provided for that position. Our results suggest that if participants engage in central processing, they are able to use both general informal-logic criteria as well as scheme-specific ones to assess the quality of the message's arguments. Which criteria will be used, may depend on several factors. If an argument contains a clearly incorrect assumption, for instance, "Paris is the capital of England", applying the (informal-logic) acceptability criterion may be easiest. Although employing argument scheme specific criteria involves the additional step of identifying the relevant argument scheme, this does not necessarily imply that it will always takes more time or cognitive energy to apply. The specific scheme can be signalled so explicitly ("I have it upon good

---

[17] Here we have collapsed the criteria *overall similarity*, *relevant similarities* and *irrelevant dissimilarities*.

[18] As indicated, we regard the criterion of *causality of example* as an artefact of the causal nature of the arguments from example in our experimental materials.

authority that ...”; “Because X says so”), that application of the relevant criteria is relatively straightforward.

When language users are centrally processing an argumentative message, arguments that satisfy criteria should be more persuasive than arguments that do so to a lesser extent and as a consequence the claims supported by these arguments should be more acceptable. In experimental studies, we found substantial support for this hypothesis: violation of the criteria *trustworthiness* (Hoeken et al. 2012, 2014) and *recency* (Hoeken et al. 2012) for argumentation from authority, *relevant similarities* and *irrelevant dissimilarities* for argumentation from analogy (Hoeken and Hustinx 2009; Hoeken et al. 2012), *number of examples* (Hoeken and Hustinx 2009; Hoeken et al. 2014; Hornikx and Hoeken 2007) and *relevance of example* for argumentation from example (Hoeken et al. 2014), *cause sufficiency* and *cause relevance* for argumentation from cause to effect (Hoeken et al. 2014) and *desirability of effect* (Hoeken et al. 2012) for argumentation from consequence resulted in a significantly lower claim acceptability. Only violation of the criteria *expertise* (for argumentation from authority) and *effectivity* (for argumentation from consequence) did not always reveal the effect predicted (Hoeken et al. 2012, 2014), possibly as a result of too subtle manipulations.

We showed in this article that—in addition to experimental research—qualitative studies into the criteria used in evaluation tasks are important to provide support for the conventional validity of normative concepts from argumentation theory and to interpret the concept of central or systematic processing in persuasion theories.

Appendix

Open case – argumentation from analogy (Timmers 2014, p. 159; translated from Dutch)

You have to write a speech for the Minister. He defends the following claim in this speech:

It is desirable that taxes on snacks high in calories (like a croquette or a Mars® bar, etc.) are increased.

As a speechwriter you are well aware that you have to support your claims with strong arguments. Your job is to come up with two arguments that support your claim. You have chosen to support your claim by means of comparisons. Provide:

One strong comparison that supports the claim well

And

One weak comparison that does not support the claim well.

After you have mentioned a strong and a weak argument, you will be asked why one argument is stronger as support for the claim than the other.

The goal of this assignment in your job interview is to find out if you are able to distinguish between strong and weak arguments. So:

It is desirable that taxes on snacks high in calories (like a croquette or a Mars® bar, etc.) are increased,

A. The strong argument:	B. The weak argument:
because that is comparable to	because that is comparable to
…………………………..	…………………………..
…………………………..	…………………………..

The question that will be asked afterwards, is: why is argument A stronger in defense of the claim than argument B?

Closed case – argumentation from analogy (Timmers 2014, p. 161; translated from Dutch)

The civil rights of prisoners are a hotly debated issue. It also receives a lot of attention from politicians. At the moment, the media pay a lot of attention to Willem van E, also known as "the monster from Harkstede". He is sentenced to life imprisonment because he murdered three prostitutes. He currently receives a AOW (financial contribution from the government) each month (as does everyone over 65).

Imagine you have to defend the following claim:

It is undesirable that Willem van E, sentenced to life imprisonment, receives AOW

To support this claim, you can choose from the arguments below. Which of these arguments do you consider strong in support of the claim? And which arguments do you consider weak in support of this claim? Rank these arguments so from 1 (strongest) to 5 (weakest).

A. A similar situation is that of John Kraaykamp Sr., 82 years old. He is left with only € 5,00 from his AOW because he has to contribute to the costs of living in the retirement home he is living in. Therefore it is undesirable that Willem van E. receives AOW.
B. You don't give candy to a naughty child, do you? Therefore, it is undesirable….
C. Pieter and Rika, two people over 65 living together, also receive less AOW because of the financial advantages of living together. Therefore, it is u ndesirable…
D. Edwin, a 40 year old construction worker, has to work for his living as well. Therefore, it is undesirable…

E. Achmed B., the murderer of Theo van Gogh, also receives no money from his disability benefit now he is in prison. Therefore, it is undesirable…

Ranking: 1. _____ 2._____ 3. _____ 4. _____5. _____

References

Blair, J.A. 2001. Walton's argumentation schemes for presumptive reasoning: A critique and development. *Argumentation* 15: 365–379.

Chaiken, S. 1987. The heuristic model of persuasion. In *Social influence: The Ontario symposium*, vol. 5, ed. M.P. Zanna, J.M. Olson, and C.P. Herman, 3–39. Hillsdale: Erlbaum.

Chen, S., and S. Chaiken. 1999. The heuristic-systematic model in its broader context. In *Dual-process theories in social psychology*, ed. S. Chaiken, and Y. Trope, 73–96. New York: Guilford.

Freeley, A.J., and L. Steinberg. 2000. *Argumentation and debate. Critical thinking for reasoned decision making*, (10th ed.) Belmont, CA: Wadsworth.

Garssen B. 1997. *Argumentatieschema's in pragma-dialectisch perspectief. Een theoretisch en empirisch onderzoek.* [Argument Schemes from a Pragma-Dialectical Perspective: A Theoretical and Empirical Study.] Proefschrift Universiteit van Amsterdam: IFOTT.

Goddon, D.M., and D. Walton. 2007. Advances in the theory of argumentation schemes and critical questions. *Argumentation* 27: 267–292.

Govier, T. 2005. *A practical study of argument*, (6th ed.) Belmont: Wadworth/Thomson Learning.

Hoeken, H., and L. Hustinx. 2009. When is statistical evidence superior to anecdotal evidence? The role of argument type. *Human Communication Research* 35: 491–510.

Hoeken, H., E. Šorm, and P.J. Schellens. 2014. Arguing about the likelihood of consequences: Laypeople's criteria to distinguish strong from weak ones. *Thinking & Reasoning* 20: 77–98.

Hoeken, H., R. Timmers, and P.J. Schellens. 2012. Arguing about desirable consequences: what constitutes a convincing argument? *Thinking & Reasoning* 18: 394–416.

Hornikx, J., and H. Hoeken. 2007. Cultural differences in the persuasiveness of evidence types and evidence quality. *Communication Monographs* 74: 443–463.

Johnson, R.H., and Blair, J.A. 2006. *Logical self-defense*. [1st edition 1977.] New York: IDEA.

Kahane, H. 1992. *Logic and contemporary rhetoric. The use of reason in everyday life*. (6th ed.) Belmont CA: Wadsworth.

Kienpointner, M. 1992. *Alltagslogik. Struktur und Funktion von Argumentationsmustern.* [Everyday Logic: Structure and Function of Argument Schemes.] Stuttgart – Bad Cannstatt: Fromann-Holzboog.

Krueger, R.A., and M.A. Casey. 2000. *Focus groups. A practical guide for applied research*, (3rd ed.) Thousand Oaks: Sage.

Meany, J., and K. Shuster. 2002. *Art, argument and advocacy. Mastering parliamentary debate*. New York: The International Debate Education Association.

Petty, R.E., and J.T. Cacioppo. 1986. *Communication and persuasion. Central and peripheral routes to attitude change*. Berlin: Springer.

Petty, R.E., and J.R. Priester. 1994. Mass media attitude change: Implications of the Elaboration Likelihood Model of persuasion. In *Media effects. Advances in theory and research*, ed. J. Bryant, and D. Zillmann, 91–122. Hillsdale: Lawrence Erlbaum.

Reinard, J. 1991. *Foundations of argument. Effective communication for critical thinking*. Dubuque: Wm. C. Brown.

Schellens, P.J. 1985. *Redelijke argumenten: een onderzoek naar normen voor kritische lezers.* [Reasonable Arguments: A Study on the Norms for Critical Readers.] Proefschrift Rijksuniversiteit Utrecht. Dordrecht: Foris Publications.

Schellens, P.J., and M. de Jong. 2004. Argumentation schemes in persuasive brochures. *Argumentation* 18: 295–323.

Šorm, E. 2010. *The good, the bad and the persuasive. Normative quality and actual persuasiveness of arguments from authority, arguments from cause to effect and arguments from example*. Proefschrift Radboud Universiteit Nijmegen. [LOT 241].

Šorm, E., R. Timmers, and P.J. Schellens. 2007. Determining laymen criteria: Evaluating methods. In *Proceedings of the sixth conference of the international society for the study of argumentation*, ed. F.H. van Eemeren, J.A. Blair, C.A. Willard, and B. Garssen, 1321–1329. Amsterdam: SicSat.

Timmers, R.H.M. 2014. *De wapens van de lezer: criteria voor argumentkwaliteit en het effect van het schenden daarvan op de overtuigingskracht van argumentatie op basis van autoriteit, analogie en voor- en nadelen.* [The Readers' Weapons: Criteria for Argument Quality and the Effect of Violating Them on the Persuasiveness of Argumentation from Authority, Analogy, and Consequences.] Proefschrift Radboud Universiteit Nijmegen.

Van Eemeren, F.H., and R. Grootendorst. 1992. *Argumentation, communication, and fallacies. A pragma-dialectical perspective*. Hillsdale: Lawrence Erlbaum.

Van Eemeren, F.H., and R. Grootendorst. 2004. *A systematic theory of argumentation. The pragma-dialectical approach*. Cambridge: Cambridge University Press.

Van Eemeren, F.H., R. Grootendorst, and A.F. Snoeck Henkemans. 2002. *Argumentation. Analysis, evaluation, presentation*. Mahwah: Lawrence Erlbaum.

Van Eemeren, F.H., B. Garssen, and B. Meuffels. 2009. *Fallacies and judgments of reasonableness. Empirical research concerning the pragma-dialectical discussion rules*. Dordrecht: Springer.

Walton, D. 1996. *Argumentation schemes for presumptive reasoning*. Mahwah, NJ: Lawrence Erlbaum.

Walton, D., C. Reed, and F. Macagno. 2008. *Argumentation schemes*. Cambridge: Cambridge University Press.

Warnick, B., and E.S. Inch. 1989. *Critical thinking and communication. The use of reason in argument*. New York: Macmillan.