Argumentation (2008) 22:555–569
DOI 10.1007/s10503-007-9067-6

Comparing the Actual and Expected Persuasiveness of Evidence Types: How Good are Lay People at Selecting Persuasive Evidence?

Jos Hornikx

Published online: 28 February 2008

Abstract Whereas there are many publications in which argumentation quality has been defined by argumentation theorists, considerably less research attention has been paid to lay people's considerations regarding argument quality. Considerations about strong and weak argumentation are relevant because they can be compared with actual persuasive success. Argumentation theorists' conceptions have to some extent been shown to be compatible with actual effectiveness, but for lay people such compatibility has yet to be determined. This study experimentally investigated lay people's expectations about the persuasiveness of anecdotal, statistical, causal, and expert evidence, and compared these expectations with the actual persuasiveness of these evidence types. Dutch and French participants ($N = 174$) ranked four types of evidence in terms of their expected persuasiveness for eight different claims. Both cultural groups expected statistical evidence to be the most persuasive type of evidence to other people, followed by expert, causal, and, finally, anecdotal evidence. A comparison of these rankings with the results of Hornikx and Hoeken (*Communication Monographs* 74, 443–463, 2007, Study 1) on the *actual* persuasiveness of the same evidence types reveals that people's expectations are generally accurate: How relatively persuasive they expect evidence types to be often corresponded with their actual persuasiveness.

Keywords Actual persuasiveness · Argument quality · Evidence · Expected persuasiveness · Perceived persuasiveness

J. Hornikx (✉)
Business Communication Studies, Centre for Language Studies,
Radboud University Nijmegen, Erasmusplein 1, P.O. Box 9103,
6500 HD Nijmegen, The Netherlands
e-mail: j.hornikx@let.ru.nl

1 Introduction

The quality of arguments has, naturally, received a lot of attention from scholars in argumentation. Argumentation theory—apart from variations in approaches—is characterized by an interest in normative criteria for strong argumentation (e.g., what arguments *should* be persuasive?). Argument quality has also been addressed in persuasive effects research (see, e.g., Petty and Cacioppo 1986). Persuasion scholars are interested in the actual effects of different arguments (e.g., what arguments *are* persuasive?). Despite this difference in perspective of argumentation theory and persuasive effects research, a confrontation of both approaches can be fruitful (O'Keefe 1995). For instance, the normative conceptions of argument quality that are distinguished in argumentation theory are highly relevant for persuasion studies in order to better determine what characterizes arguments that have been found to be strong and arguments that have been found to be weak.

Furthermore, it is insightful to compare normative and descriptive perspectives on argument quality. In a number of publications, O'Keefe has compared normatively-desirable argumentative practice with practical persuasive success (O'Keefe 2002b, 2005, 2007). A recent overview of research (O'Keefe 2005) showed some support for compatibility between what argumentation theorists think *should* be persuasive argumentation, and what *is* persuasive in practice. To give one example, messages that contain an explicit statement of the conclusion, which is normatively desirable (Van Eemeren et al. 1993, p. 173), are generally more persuasive than messages that omit such a statement (O'Keefe 2002b).

Whereas argumentation theorists have extensively documented their conceptions of what constitutes normatively high and low argument quality (see, e.g., Garssen 1997; Schellens 1985; Van Eemeren et al. 1996), considerably less research attention has been paid to lay people's conceptions of argument quality. These conceptions are likely to differ from scholars' conceptions, as scholars simply have more knowledge of the issues at hand (for a discussion, see O'Keefe 1993; for empirical evidence, see Murphy 2001). One consequence is that lay people's conceptions of strong argumentation may be less accurate than those of scholars. A question that naturally arises here, and that has hardly been addressed in previous research, is: How good are the conceptions of persuasive arguments that normal language users have?[1] A comparison of lay people's expectations about strong arguments and the actual persuasiveness of arguments provides an answer to this question: Are the arguments that people expect to be relatively strong also relatively persuasive in practice?

I will report on an experiment that was conducted to answer this question, focusing on the effects of evidence inclusion on claim acceptance. In Sect. 3, I will introduce the concept of evidence, discuss how evidence is related to persuasive argumentation, and present a typology of different evidence types. The question about people's accuracy at selecting persuasive evidence relates to knowledge about

[1] An example of a study that investigated lay people's accuracy at evaluating normatively strong argumentation is Jacobs et al. (1985). In two studies, it was demonstrated that lay people prefer valid conclusions from syllogisms to invalid conclusions when these valid conclusions were argued for by another person.

the expected persuasiveness of evidence types, and about the actual persuasiveness of evidence types. Therefore, a short review of relevant work on the expected and actual persuasiveness of evidence types will be provided (Sect. 4). I will, however, start with a general discussion about expected persuasiveness, because it can easily be mistaken with the seemingly corresponding notion of *perceived* persuasiveness (Sect. 2).

2 Actual, Perceived, and Expected Persuasiveness

Broadly speaking, persuasive effects research is interested in the effects of source factors (e.g., credibility), message factors (e.g., evidence), receiver factors (e.g., involvement), and situational factors (e.g., distraction) on persuasiveness. The perspective of this research is often receiver-oriented: What are the effects of such factors on receivers? Empirical studies in this field examine what the actual persuasive effects are of, for example, the source's credibility on receivers' attitude. Whereas actual persuasiveness is about what *is* persuasive, perceived persuasiveness is about what is *thought* to be persuasive. More specifically, perceived persuasiveness has been used to capture the notion of people's thoughts about how much a particular message persuaded them. After reading a message, people may be asked to respond to questions such as "How much to you believe your opinions on this issue were influenced by this message?". Parrott et al. (2005, p. 440), for instance, asked their participants to fill in three semantic differential scales (persuasive—not persuasive, convincing—not convincing, useful—not useful) following the stem "I think the information in the message is".

Dillard et al. (2007) meta-analytically examined studies that included measures of perceived and actual persuasiveness. Their analysis shows a positive and substantial relationship between perceived and actual persuasiveness. This means there is some consistency between what people think persuaded them and what actually did persuade them (but see O'Keefe 1993, 2002a). Lay people, however, do not always correctly assume do not necessarily know how much an argument or a credible source persuaded them. As O'Keefe (1993, p. 231) explained, such people "will naturally base their answers on their (implicit or explicit) beliefs about what persuades. But because these beliefs may be erroneous, respondents' judgments of the likelihood of influence may not correspond with the actual likelihood of influence". In Hoeken (2001), the perceived persuasiveness of anecdotal, statistical, and causal evidence was assessed by having participants rate how sound, relevant, strong, and convincing they found one of the three types of evidence they had read in the message. These assessments demonstrated that statistical evidence was perceived to be more persuasive than anecdotal evidence. The actual persuasiveness of these evidence types, however, was the same: anecdotal and statistical evidence were equally successful in convincing participants of the claim in the message.

Perceived persuasiveness should not be confounded with *expected* persuasiveness. Both kinds of persuasiveness are about people's conceptions about whether or not a message or an argument is persuasive, but they differ in the people at which this message is targeted. Perceived persuasiveness is receiver-oriented, and is

directed to the person who assesses the perceived persuasiveness. In fact, it is a message recipient's judgment about how much a message (or argument) persuaded him or her. Expected persuasiveness could be said to be sender-oriented: How persuasive does a message sender (e.g., a participant in a study) think an argument is for *another* person. The conceptions of argumentation theorists about argument quality are to be regarded as conceptions of expected persuasiveness, because these conceptions are (well-reasoned) thoughts about how strong arguments should be for others. It is specifically this notion of expected persuasiveness that is needed when faced with the question as to how good lay people are at selecting persuasive arguments. Such expected persuasiveness should then be compared with actual persuasiveness. In this paper, the comparison of expected and actual persuasiveness involves types of evidence as support for claims. Below, the concept of evidence and the types of evidence will be introduced.

3 The Role of Evidence in Persuasive Argumentation

Persuasive messages seek to influence people's attitudes towards an object or behavior. Very often, such messages can be considered as texts involving pragmatic argumentation. In the scheme of pragmatic argumentation, behavior is recommended on the basis of its favorable consequence(s) (positive variant) or advised against on the basis of its unfavorable consequence(s) (negative variant) (Feteris 2002). An example of the positive variant is the advice to students to listen to classical music (behavior), because this helps them to absorb a lot of knowledge in a short period of time (favorable consequence). The simplest form of this positive variant of pragmatic argumentation looks like (1):

(1) behavior A leads to consequence B
consequence B is desirable
therefore: behavior A is desirable

According to this reasoning, listening to classical music is desirable because absorbing knowledge in an efficient way is desirable. Argument theorists claim that at least two questions should come to readers' minds when they assess such a line of argumentation (e.g., Feteris 2002; Garssen 1997; Schellens 1985). The first question is about the desirability of the consequence: Is absorbing knowledge in an efficient way desirable? The second question is about the probability of the consequence: Will listening to classical music indeed help students to absorb a lot of knowledge in a short period of time? A pragmatic argument is normatively strong if these critical questions can be answered affirmatively. A few studies have demonstrated that it is much easier for people to assess a consequence's desirability than to assess its probability (Areni and Lutz 1988; Van Enschot-Van Dijk et al. 2003). As a result, the consequence's probability may need supplementary support. As a matter of fact, a corpus study has shown that text writers more frequently use additional information to support claims about the probability of consequences than claims about their desirability (Hornikx et al. 2003).

Table 1 Four types of evidence for the claim 'Listening to classical music helps students to absorb a lot of knowledge in a short period of time'

Anecdotal evidence	The 16-year-old Martijn Mulder from The Hague has been able to absorb a lot of knowledge in a short period of time, since he regularly listened to classical music.
Causal evidence	Classical music stimulates the recognition of repeating patterns and complicated structures, so that analytic thought is developed, and a lot of knowledge can be absorbed.
Expert evidence	Professor Dr. Wildschut, a specialist in the field of music studies at the University of Maastricht, underscores that students can absorb a lot of knowledge in a short period of time by listening to classical music.
Statistical evidence	The results of a study among 322 Dutch students showed that 75% of them had absorbed a lot of knowledge in a short period of time by listening to classical music.

In persuasion studies, this additional information is usually called evidence. Evidence is defined as "data (facts or opinions) presented as proof for an assertion" (Reynolds and Reynolds 2002, p. 429). Hoeken and Hustinx (2003) distinguish four types of evidence: anecdotal, statistical, causal, and expert evidence.[2] Anecdotal evidence consists of one case, whereas statistical evidence consists of numerical information about a large number of cases. Causal evidence consists of an explanation, and expert evidence, finally, consists of a confirmation by an expert. Table 1 gives examples of the four evidence types for the claim about the positive effect of listening to classical music.

4 Studies on the Actual and Expected Persuasiveness of Evidence Types

A considerable number of experimental studies have examined the actual persuasive effects of evidence types in different fields, such as advertising, educational communication, health communication, and argumentation. In particular, the persuasiveness of anecdotal and statistical evidence has been extensively studied (see reviews in Allen and Preiss 1997; Baesler and Burgoon 1994; Reinard 1988). Whereas earlier reviews concluded that anecdotal evidence was more persuasive than statistical evidence, the statistical summary that Allen and Preiss (1997) provided in their meta-analysis concluded that it is statistical evidence that is the most persuasive of the two evidence types (for an explanation of these inconsistent conclusions, see Allen et al. 2006; Hornikx 2007). Hornikx (2005) was the first review to include studies in which two or more evidence types were compared. In this narrative review, statistical and causal evidence were found to be more persuasive than anecdotal evidence, and expert evidence tended to be more persuasive than anecdotal evidence. In contrast to the tradition of empirical studies on the actual persuasiveness of evidence types, there are only a few studies that have examined their expected persuasiveness. These studies are presented below.

[2] As Hornikx and Hoeken (2007) explain, there is a relation between evidence types and argument types. Whereas evidence focuses on the type of data that support claims, argument types focus on the relationship between the argument (evidence) and the claim. There is no one-to-one relationship between evidence types and argument types, because anecdotal evidence and statistical evidence are both related to arguments by generalization.

The relative frequency with which evidence is used in argumentative discourse can be taken as an indication for people's expectations of strong and weak evidence. The implicit reasoning here is that the kinds of evidence people use most are also the kinds of evidence they expect to be most effective. In an argument production study, Kline (1971b), for instance, investigated the use of evidence types in students' speeches. He hypothesized that highly dogmatic students would use more documented evidence attributed to sources than undocumented evidence. Dogmatic people are said to be closed-minded, intolerant, and deferential to authority (Rokeach 1960). Students were given a lecture in which various types of evidence were presented. About 2 weeks later, each student gave a short speech. The students' use of documented and undocumented evidence in these speeches was analyzed. Highly dogmatic participants were indeed found to use more documented evidence than participants who scored low on the dogmatism scale.

Levasseur and Dean (1996) analyzed the actual use of evidence types in debates between American presidential candidates. They distinguished anecdotal (specific historical instances), statistical (numerical instances), and source evidence (relying on experts or lay people). Results showed that the presidential candidates used statistical and anecdotal evidence about four times more than source evidence. Hornikx et al. (2003), finally, analyzed the relative frequency with which Dutch and French text writers of persuasive information leaflets use anecdotal, statistical, causal, and expert evidence. In these leaflets—covering topics such as cancer, the Third World, and alcohol consumption—anecdotal, causal, and statistical evidence were more frequently used than expert evidence. These results were, however, qualified by the cultural background of the writers. Causal evidence was more frequent in the Dutch corpus, and statistical and expert evidence were more frequent in the French corpus.

The three studies discussed above show that people use some types of evidence more frequently than others in speeches or texts. A limitation of these studies is that the writers or speakers may not always have had all types of evidence at their disposal when they wanted to use evidence. When senders can explicitly choose between different evidence types provided, expected persuasiveness can be more appropriately measured. Such a design was used in an argument production study that was probably the first to examine the expected persuasiveness of evidence types (Kline 1971a). Participants were given a claim, either 'Requirements for admission to college should be raised', or 'The Federal government should exercise more control over primary and secondary education'. The claim was followed by 25 instantiations of evidence that differed on the dimensions of specificity and relevance and on the attribution to different sources. Participants ranked these instantiations according to the likelihood that they would use them to persuade either college students of the college admission claim, or the local Parents–Teachers Association of the government control claim. Kline's main goal was to investigate whether people differed in their rankings. Response patterns were analyzed in such a way that participants were clustered in different groups of encoders. Kline showed that people can indeed differ in their expectations of how persuasive different evidence types are to other people.

5 Research Questions

Whereas the actual persuasiveness of anecdotal, statistical, and—to a lesser degree—causal, and expert evidence has been extensively studied, their expected persuasiveness has never been examined using a ranking method that Kline (1971a) employed. Kline (1971a, p. 190) argued that studying the expected persuasiveness of evidence is an important addition to studies on actual persuasiveness:

> "Such studies of the [actual persuasive] effects of evidence on audiences are useful, but they are not sufficient for an adequate understanding of communication, since an audience responding to a message is only one of the important parts of the process. Among other things, we must also study the encoding part, e.g., the way in which sources select materials for their messages"

Others, such as O'Keefe (2002a) and Wilson (2002), have underscored the interest of the study of persuasion from both a sender (expected persuasiveness) and a receiver (actual persuasiveness) perspective. The first research question therefore reads:

RQ1 How persuasive do people expect anecdotal, statistical, causal, and expert evidence to be for others?

Argumentation theorists' conceptions of desirable argumentation have to some extent been shown to be compatible with actual argument effectiveness in practice (e.g., O'Keefe 2002b, 2005), but for lay people such compatibility has yet to be determined. Therefore, the question is addressed as to how good the conceptions of persuasive evidence are that normal language users have:

RQ2 How good are lay people at selecting persuasive evidence?

6 Method

The expected persuasiveness of the four evidence types (RQ1) was investigated experimentally with the ranking method in which multiple types of evidence are available to message designers (cf. Kline 1971a). An argument production study was conducted in which participants were given a series of claims, each followed by four types of evidence. The participants were asked to rank these evidence types according to their expected persuasiveness to convince another person. Actual persuasiveness was not investigated in this study because of the risk of interference between participants' expectations and their actual appreciation of evidence types. People's accuracy at selecting persuasive evidence (RQ2) was examined by comparing the current experiment's results with those of Hornikx and Hoeken (2007, Study 1), a recent study on the actual persuasiveness of the same evidence types. In Hornikx and Hoeken (2007), two groups of participants were distinguished: Dutch and French students. The same two groups were used in the current experiment in order to render the study's investigation of expected persuasiveness and of lay people's conceptions more robust.

6.1 Material

From the 20 claims used in Hornikx and Hoeken (2007), eight claims were randomly selected. As these claims described the occurrence of effects as a result of an action, they could be classified as descriptive, causal claims. Topics ranged from school performance and work productivity to party games and cartoons. The claims had been pretested to be moderately probable (Hornikx and Hoeken 2007). For each of these claims, a set of four types of evidence was constructed. An example of the operationalizations of the four types of evidence, borrowed from Hornikx and Hoeken (2007), is shown in Table 1.

As suggested by Reinard (1988) and Reynolds and Reynolds (2002), the quality of evidence was taken into account. Normatively strong instantiations of statistical and expert evidence were created on the basis of normative criteria from argumentation theory (see Van Eemeren et al. 1996; Garssen 1997). Normatively strong statistical evidence should consist of a large sample of cases that is representative for the population in the claim that it supports. The statistical evidence instantiations therefore cited studies that had a large sample size (e.g., 322 people), and contained a high percentage of cases in the sample for which the claim held (e.g., 77%). Expert evidence is normatively strong if the expert is credible and reliable, and if the expert's field of expertise corresponds to the field of the claim. Credibility and reliability were operationalized together through the titles of the expert (Prof. Dr.), and through relevant fields of expertise related to the claims under consideration (e.g., music studies in the case of the claim in Table 1). Anecdotal evidence is based on only one case supporting the general claim about a whole population; as a consequence, a normatively strong operationalization was impossible. Anecdotal evidence consisted of a short sentence about a person who had experienced the consequence (e.g., absorbing knowledge in an efficient way) as a result of behavior (e.g., listening to classical music). Causal evidence should provide a mechanism that explains the causal relation between the cause and the effect in the claim. As it is impossible to determine a normatively strong mechanism, causal evidence consisted either of two causal relations (which is typical of everyday explanations; Hesslow 1988), or of a causal chain. The Dutch claims and evidence were translated into French by native Dutch speakers, and then back translated into Dutch by native French speakers (Hornikx and Hoeken 2007). Equivalence was further increased through a careful selection of Dutch and French first names, last names, places, and universities.

6.2 Participants

In correspondence with Hornikx and Hoeken (2007), the participants were students. The Dutch participants ($n = 88$) studied Business Communication Studies at the Radboud University Nijmegen. The French participants studied Applied Foreign Languages (University of Montpellier-II, $n = 56$) or Linguistics (University of Paris-VIII, $n = 30$). Questionnaires filled in by non-natives were discarded from analyses. The percentage of male participants was higher in France (45.3%) than in

the Netherlands (14.8%) ($\chi^2(1) = 19.41$, $p < 0.001$). On average, the Dutch participants were 19.48 ($SD = 1.95$) years old, and the French participants 22.05 ($SD = 2.22$). This difference was significant: $t(172) = 8.12$, $p < 0.001$. Participants' gender and age hardly affected the expected persuasiveness of the types of evidence.[3]

6.3 Design

Four versions of the material were created. In the versions, the order of the eight claims was identical, but the order of presentation of the types of evidence for each of the claims was different. A balanced Latin square design was employed to assign the different orders of the types of evidence to the eight claims, and to the four versions.[4]

6.4 Instrumentation

The booklet that participants received was titled 'Convincing another person with arguments'. In a written instruction, participants were asked to rank the four types of evidence in terms of their persuasive power to convince another person of the believability of each of the eight claims. They indicated their ranking by noting down a '1' (expected to be most persuasive), '2', '3', or '4' (expected to be least persuasive) for each evidence type. For each claim, each number had to be used, but it could only be used once. The questionnaire ended with questions about participants' age, gender, nationality, and current education.[5]

[3] Participants' gender did not affect the mean rankings of statistical, causal, or expert evidence ($ps > 0.10$), but it did have an effect on anecdotal evidence ($z = 3.78$, $p < 0.001$). Men ($M = 3.32$, $SD = 0.75$) ranked anecdotal evidence as more persuasive than women ($M = 3.67$, $SD = 0.53$). Next, age only significantly correlated with statistical evidence: $r(174) = -0.21$, $p < 0.01$ (other evidence types: $ps > 0.05$).

[4] The rankings were pooled across the four versions, because there were no effects of version on the mean ranking of anecdotal evidence (Kruskal–Wallis $\chi^2(3) = 6.12$, $p = 0.11$), statistical evidence ($\chi^2(3) = 3.81$, $p = 0.28$), causal evidence ($\chi^2(3) = 1.05$, $p = 0.79$), or expert evidence ($\chi^2(3) = 5.15$, $p = 0.16$).

[5] As a check with Hornikx and Hoeken (2007, Study 1), this questionnaire also included measures that were relevant to the cross-cultural investigation of expert evidence in that study. After the eight rankings of the evidence types, the Preference for Expert Information scale (PEI; Hornikx and Hoeken 2007), and seven items of the Need for Cognition scale (NFC; Cacioppo et al. 1984) were included. The first four items of the PEI scale proved to be reliable, both for the French participants ($\alpha = .73$), and for the Dutch participants ($\alpha = .75$). Dutch ($M = 2.51$, $SD = 0.69$) and French participants ($M = 11\ 2.52$, $SD = 0.77$) scored equally on the PEI scale ($t(172) = 0.95$, $p = .35$). The NFC scale was reliable for the French participants ($\alpha = .74$), but not for the Dutch ($\alpha = .58$). For each expert in the expert evidence, participants indicated their perceived expertise. The experts were considered as persons with relatively high expertise, as the mean perceived expertise scored above the midpoint (3.00) of the scale for the French ($M = 3.25$, $SD = 0.60$; $t(85) = 3.82$, $p < .001$) and the Dutch participants ($M = 3.52$, $SD = 0.54$; $t(87) = 9.17$, $p < .001$).

Table 2 Mean rankings of expected persuasiveness

Evidence type	Dutch participants ($n = 88$)		French participants ($n = 86$)	
	M	SD	*M*	SD
Statistical	1.25[a]	0.44	1.55[a]	0.67
Expert	2.33[b]	0.62	2.33[b]	0.69
Causal	2.68[c]	0.67	2.71[c]	0.78
Anecdotal	3.73[d]	0.49	3.40[d]	0.70

Note: A smaller number indicates higher expected persuasiveness; different superscripts in the same column refer to significant differences between the means, alpha level of .001

6.5 Procedure and Statistical Tests

The questionnaires were filled in at the three afore-mentioned universities. The questionnaires were distributed at the beginning of a lecture. After the questionnaires had been collected, the real research purpose was revealed, and participants were thanked for their cooperation. The students received no reward for their participation. The procedure took about 15 min. A Friedman test was used to investigate whether the mean rankings of the four types of evidence differed. Next, a Wilcoxon signed rank test was employed to test which types of evidences differed from each other in their mean ranking.

7 Results

The first research question that was of interest here is: How persuasive do people expect anecdotal, statistical, causal, and expert evidence to be for others? Table 2 provides a ranking of the types of evidence from highest (1.00) to lowest (4.00) expected persuasiveness.

As Table 2 clearly shows, there was a difference in the mean rankings of the four types of evidence (Friedman $\chi^2(3) = 284.63$, $p < 0.001$). The mean rankings were compared for each culture separately. Both the Dutch and the French participants expected statistical evidence to be most persuasive, followed by expert, causal, and anecdotal evidence. This result was found to be robust, since this pattern proved to hold across the eight claims. For each individual claim, the ranking of the four types of evidence was similar to the ranking averaged over all claims.[6]

The second question that this study addressed was: How good are lay people at selecting persuasive evidence? This question is answered by comparing the expected persuasiveness of evidence types (Table 2) with the actual persuasiveness of the same evidence types found in Hornikx and Hoeken (2007, Study 1). The

[6] The range of rank scores for the eight claims for the Dutch was: 1.19–1.31 for statistical, 2.24–2.39 for expert, 2.56–2.87 for causal, and 3.67–3.78 for anecdotal evidence. For the French, the range of rank scores was: 1.36–1.71 for statistical, 2.22–2.47 for expert, 2.55–2.92 for causal, and 3.27–3.51 for anecdotal evidence.

Table 3 A comparison of the rankings of expected and actual persuasiveness

Evidence type	Dutch participants		French participants	
	$n = 88$ Expected	$n = 305$ Actual	$n = 86$ Expected	$n = 295$ Actual
Statistical	1	1	1	1
Expert	2	2	2	1
Causal	3	2	3	1
Anecdotal	4	4	4	4

Note: A smaller number indicates higher persuasiveness. Actual persuasiveness is based on results found in Hornikx and Hoeken (2007, Study 1)

participants in the two studies were highly comparable.[7] How does the expected persuasiveness of evidence types relate to the actual persuasiveness of evidence types? Table 3 compares the rankings of the four types of evidence in this study with the rankings that were made on the basis of the results of Hornikx and Hoeken (2007, Study 1) on actual persuasiveness.

In the columns about actual persuasiveness, different numbers represent means that significantly differed. For the Dutch participants, for instance, statistical evidence was more persuasive than expert and causal evidence, which were equally persuasive.

The conceptions that the students participating in the current study had about the persuasiveness of evidence types are accurate, if the rankings of actual and expected persuasiveness are highly similar. The participants, in particular the Dutch appeared to be generally good at selecting persuasive evidence for others: The types of evidence they expected (not) to be persuasive, were in most cases actually (not) persuasive (see Table 3). For both cultural groups, anecdotal evidence was expected to be least persuasive, which corresponds with its low actual persuasiveness. Of the eight comparisons between expected and actual persuasiveness, there are three differences between the rankings. For the Dutch group of participants, there only was a difference between expected and actual persuasiveness for causal evidence. That is, these participants expected expert evidence to be more persuasive than causal evidence, but the actual persuasiveness of these evidence types did not differ in Hornikx and Hoeken (2007, Study 1). There were more differences for the French group of participants. French participants expected statistical evidence to be more persuasive than expert evidence, and expert evidence to be more persuasive than causal evidence, but in reality it appears that the three types of evidence are equally persuasive. The differences in the actual persuasiveness of these three types of evidence were not significant, but the pattern is somewhat similar to that of expected

[7] In both studies, the participants were mostly students at Humanities faculties. The Dutch students in this study ($M = 19.48$) were younger than in Hornikx and Hoeken (2007, Study 1) ($M = 20.98$), and the French students in this study ($M = 22.05$) were older than in Hornikx and Hoeken (2007, Study 1) ($M = 20.75$). In this study, the percentage of male students was 45.3% (French) or 14.8% (Dutch), whereas the percentages were lower in Hornikx and Hoeken (2007, Study 1) (Dutch: 22.6%; French: 13.2%). In both studies, however, age and gender hardly affected the actual or expected persuasiveness of evidence types.

persuasiveness: expert evidence, followed by statistical evidence, followed by causal evidence (Hornikx and Hoeken 2007, Study 1).

8 Conclusion and Discussion

The conceptions of normatively desirable argumentation that argumentation theorists have formulated have to some extent been found to be compatible with actual effects that argumentation has in practice (for an overview, see O'Keefe 2005). As such compatibility had yet to be determined for lay people, the current study was conducted to provide insights into lay people's accuracy at designing high quality argumentation. In particular, the expected persuasiveness of anecdotal, statistical, causal, and expert evidence was investigated, and compared with the actual persuasiveness of the same evidence types found in Hornikx and Hoeken (2007, Study 1). Effort was put in creating a high degree of similarity between the two studies, such as in the kind of participants, the evidence operationalizations, and the claims involved.

The results can be summarized as follows. First of all, both Dutch and French participants expected statistical evidence to be the most persuasive type of evidence to other people, followed by expert, causal, and, finally, anecdotal evidence (RQ1). A comparison of these rankings with the actual persuasive effects of evidence types in Hornikx and Hoeken (2007, Study 1) reveals that people's expectations are generally accurate: How relatively persuasive they expect evidence types to be often corresponded with their actual persuasiveness (RQ2). There were a few differences between the rankings of the four types of evidence in terms of their expected and actual persuasiveness. It is plausible that some of them can be attributed to the difference in the research designs of the two studies. In the study of Hornikx and Hoeken (2007) on actual persuasiveness, participants judged the probability of a number of claims supported by evidence on rating scales. These probability scores led to a pattern of results for the actual persuasiveness of the types of evidence. Some types of evidence were found to be equally persuasive; for the French participants, in particular, expert, statistical, and causal evidence were equally persuasive. In the current experiment, however, participants were not allowed to give evidence types an equal rank, even if they expected two or three types of evidence to be equally persuasive for others. The French participants, to continue the example, may have expected expert, statistical, and causal evidence to be equally persuasive for others, but they were not able to express this expectation.

This characteristic of the research design is one of limitations of the current study. A drawback of the ranking method (cf. Kline 1971a) is that differences in the relative expected persuasiveness of two types of evidence may sometimes be artificial. Future research may explore the use of rating scales, where participants use scales from 'very persuasive for others' to 'not very persuasive for others' for each type of evidence individually. Another limitation of this study was the vagueness of 'another person' for which participants had to rank evidence types. Participants were free to think about others as targets of the claims with evidence, so their thoughts may have gone to other students, their family, or their best friends.

Indeed, it is plausible that the specific target matters when it comes to judging the effects of claims on target receivers (cf. White 1997). Therefore, more precise target receivers for each of the eight claims would have been a better option, as was done in Kline (1971b)—sidestepping the possible difficulties of determining suitable target receivers for each of the claims that were used.

This study makes three contributions to our understanding of normative and descriptive considerations of argument quality. First, the experiment presented in this paper—the first to have used ranking technique to investigate the expected persuasiveness of four types of evidence occuring in persuasive discourse (Hornikx et al. 2003)—demonstrates that Dutch and French students expect statistical evidence to be most effective for others, followed by expert, causal, and anecdotal evidence. Second, the comparison of the current experiment with Hornikx and Hoeken (2007, Study 1) contributes to our understanding of lay people's accuracy at selecting persuasive evidence for others. In fact, the conceptions of lay people from two different cultures were found to be rather good. Third, lay people's conceptions seem to be somewhat similar to experts' conceptions. A first piece of evidence for this suggestion is that statistical evidence, which is normatively stronger than anecdotal evidence because the generalization is based on numerous cases instead of on one case (e.g., Garssen 1997), was expected to be most persuasive for others, whereas anecdotal evidence was expected to be least persuasive. A second piece of evidence is related to the finding that the two types of evidence that were operationalized as normatively strong were also expected to be most persuasive by lay people: statistical and expert evidence. This result adds some support to the compatibility of lay people's considerations of argument quality on the one hand, and normative conceptions of argument quality on the other.

Further studies could investigate lay people's expectations of argument quality in other domains than that of evidence types, such as argumentation schemes and fallacies. When it comes to evidence, one avenue for future research could be the investigation of lay people's motivations behind their expectations. Why, for instance, do they believe expert evidence to be more persuasive than causal evidence? Another avenue is to study lay people's expectations of normatively strong and normatively weak evidence (cf. Hornikx and Hoeken 2007, Study 2). How do people rank evidence that is operationalized as normatively weak? In addressing such questions, further research could raise our understanding of the expected and actual persuasiveness of evidence types, and of lay people's accuracy at selecting persuasive evidence.

Acknowledgements Thanks to Chantal Claudel and Sonia Gouirand for their practical help with the experiment, and to Hans Hoeken and Daniel O'Keefe for discussions of issues addressed in this article.

References

Allen, M., and R.W. Preiss. 1997. Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. *Communication Research Reports* 14: 125–131.

Allen, M., R.W. Preiss, and B.M. Gayle. 2006. Meta-analytic examination of the base-rate fallacy. *Communication Research Reports* 23, 45–51.

Areni, C.S., and R.J. Lutz. 1988. The role of argument quality in the elaboration likelihood model. In *Advances in consumer research*, ed. M.J. Houston, vol. 15, 197–203. Provo: Association for Consumer Research.

Baesler, E.J., and J.K. Burgoon. 1994. The temporal effects of story and statistical evidence on belief change. *Communication Research* 21: 582–602.

Cacioppo, J.T., R.E. Petty, and C.F. Kao. 1984. The efficient assessment of need for cognition. *Journal of Personality Assessment* 48, 306–307.

Dillard, J.P., K.M. Weber, and R.G. Vail. 2007. The relationship between the perceived and actual effectiveness of persuasive messages: A meta-analysis with implications for formative campaign research. *Journal of Communication* 57, 613–631.

Feteris, E.T. 2002. Pragmatic argumentation in a legal context. In *Advances in pragma-dialectics*, ed. F.H. van Eemeren, 243–260. Amsterdam/Newport News: Sic Sat/Vale Press.

Garssen, B.J. 1997. *Argumentatieschema's in pragma-dialectisch perspectief: Een theoretisch en empirisch onderzoek [Argumentation schemes in a pragma-dialectical perspective: A theoretical and empirical study]*. Amsterdam: IFOTT.

Hesslow, G. 1988. The problem of causal selection. In *Contemporary science and natural explanation: Commensense conceptions of causality*, ed. D.J. Hilton, 11–32. Brighton: Harvester Press.

Hoeken, H. 2001. Anecdotal, statistical, and causal evidence: Their perceived and actual persuasiveness. *Argumentation* 15: 425–437.

Hoeken, H., and L. Hustinx. 2003. The relative persuasiveness of different types of evidence. In *Proceedings of the 5th conference of the International Society for the Study of Argumentation*, eds. F.H. van Eemeren, J.A. Blair, C.A. Willard, and A.F. Snoeck Henkemans, 497–501. Amsterdam: Sic Sat.

Hornikx, J. 2005. A review of experimental research on the relative persuasiveness of anecdotal, statistical, causal, and expert evidence. *Studies in Communication Sciences* 5: 205–216.

Hornikx, J. 2007. Is anecdotal evidence more persuasive than statistical evidence? A comment on classic cognitive psychological studies. *Studies in Communication Sciences* 7: 151–164.

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

Hornikx, J., M. Starren, and H. Hoeken. 2003. Cultural influence on the relative occurrence of evidence types. In *Proceedings of the 5th conference of the International Society for the Study of Argumentation*, eds. F.H. van Eemeren, J.A. Blair, C.A. Willard, and A.F. Snoeck Henkemans, 531–536. Amsterdam: Sic Sat.

Jacobs, S., M. Allen, S. Jackson, and D. Petrel. 1985. Can ordinary actors recognize a valid conclusion if it walks up and bites them on the butt? In *Argument and social practice: Proceedings of the 4th SCA/AFA conference on argumentation*, eds. J.R. Cox, M.O. Sillars, and G.B. Walker, 665–674. Annandale: Speech Communication Association.

Kline, J.A. 1971a. A Q-analysis of encoding behavior in the selection of evidence. *Speech Monographs* 37: 190–197.

Kline, J.A. 1971b. Dogmatism of the speaker and selection of evidence. *Speech Monographs* 38: 354–355.

Levasseur, D., and K.W. Dean. 1996. The use of evidence in presidential debates: A study of evidence levels and types from 1960 to 1988. *Argumentation and Advocacy* 32: 129–142.

Murphy, P.K. 2001. What makes a text persuasive? Comparing students' and experts' conceptions of persuasiveness. *International Journal of Educational Research* 35: 675–698.

O'Keefe, D.J. 1993. Understanding social influence: Relations between lay and technical perspectives. *Communication Studies* 44: 228–238.

O'Keefe, D.J. 1995. Argumentation studies and dual-process models of persuasion. In *Perspectives and Approaches: Proceedings of the 3*rd *ISSA Conference on Argumentation*, eds. F.H. van Eemeren, R. Grootendorst, J.A. Blair, and C.A. Willard, 3–17. Amsterdam: Sic Sat.

O'Keefe, D.J. 2002a. *Persuasion: Theory and research*. 2nd ed. Thousand Oaks: Sage.

O'Keefe, D.J. 2002b. The persuasive effects of variation in standpoint articulation. In *Advances in Pragma-dialectics*, ed. F.H. van Eemeren, 65–82. Amsterdam/Newport News: Sic Sat/Vale Press.

O'Keefe, D.J. 2005. News for argumentation from persuasion effects research: Two cheers for reasoned discourse. In *Critical problems in argumentation: Selected papers from the 13th biennial conference on argumentation*, ed. C.A. Willard, 215–221. Washington: National Communication Association.

O'Keefe, D.J. 2007. Potential conflicts between normatively-responsible advocacy and successful social influence: Evidence from persuasion effects research. *Argumentation* 21: 151–163.

Parrott, R., K. Silk, K. Dorgan, C. Condit, and T. Harris. 2005. Risk Comprehension and judgments of statistical evidentiary appeals: When a picture is not worth a thousand words. *Human Communication Research* 31: 423–452.

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

Reinard, J.C. 1988. The empirical study of the persuasive effects of evidence: The status after fifty years of research. *Human Communication Research* 15: 3–59.

Reynolds, R.A., and J.L. Reynolds. 2002. Evidence. In *The persuasion handbook: Developments in theory and practice*, eds. J.P. Dillard, and M. Pfau, 427–444. Thousand Oaks: Sage.

Rokeach, M. 1960. *The open and closed mind*. New York: Basic Books.

Schellens, P.J. 1985. *Redelijke argumenten: Een onderzoek naar normen voor kritische lezers [Reasonable arguments: A study in criteria for critical readers]*. Dordrecht: Foris.

Van Eemeren, F.H., R. Grootendorst, S. Jackson, and S. Jacobs. 1993. *Reconstructing argumentative discourse*. Tuscaloosa: University of Alabama Press.

Van Eemeren, F.H., R. Grootendorst, and F. Snoeck Henkemans. 1996. *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments*. Mahwah: Lawrence Erlbaum.

Van Enschot-Van Dijk, R., L. Hustinx, and H. Hoeken. 2003. The concept of argument quality in the elaboration likelihood model: A normative and empirical approach to Petty and Cacioppo's 'strong' and 'weak' arguments. In *Anyone who has a view: Theoretical contributions to the study of argumentation*, eds. F.H. van Eemeren, J.A. Blair, C.A. Willard, and A.F. Snoeck Henkemans, 319–335. Dordrecht: Kluwer.

White, H.A. 1997. Considering interacting factors in the third-person effect: Argument strength and social distance. *Journalism and Mass Communication Quarterly* 74: 557–564.

Wilson, S.R. 2002. *Seeking and resisting compliance: Why people say what they do when trying to influence others*. Thousand Oaks: Sage.