

ISSA Proceedings 2002 - The Relative Persuasiveness Of Anecdotal, Statistical, Causal And Expert Evidence



1. Introduction

Argument quality plays an important role in popular models of the persuasion process, such as the Elaboration Likelihood Model (ELM; Petty & Cacioppo, 1986) and the Heuristic Systematic Model (HSM; Chaiken, 1987). According to these models, argument quality determines the outcome of the persuasion process if people are motivated and able to scrutinize the message. McGuire (2000) notes that given the importance of argument quality for persuasion, remarkably little research has been done that addresses the question whether the strength of arguments is related to the type of warrant or type of evidence that is used. In this paper, we try to contribute to answering this question.

1.1 Arguments in persuasive documents: the importance of probability and desirability

Persuasive documents are often designed to influence people's behavior: to quit smoking, to drink less, or to buy a Volkswagen beetle. The usually implicit claim is that the propagated behavior, for instance, quitting smoking, is better than the alternative behavior (continuing to smoke, drink heavily, or buy another make of car). Claims about the (relative) favorableness of an option are often backed up by pragmatic argumentation (van Eemeren, Grootendorst, Snoeck Henkemans et al., 1996, 111-112). In this type of argumentation, the claim is supported by referring to the favorable consequences that acceptance of the claim will have. For instance, as a result of no longer smoking, it is argued that your chances of dying of lung cancer or of a heart attack are reduced, your physical condition improves, and your breath smells fresher.

Crucial to the persuasiveness of this type of argumentation is that the audience accepts the desirability of the consequences and that it regards it as plausible that the propagated behavior will indeed lead to these consequences. A strong

argument in favor of such a claim would be that the behavior very probably will result in a very desirable consequence. A strong argument against such a claim would be that the behavior will very probably result in a very undesirable consequence. If the audience does not readily accept the consequence's probability or desirability, you have to provide evidence to back up your claim.

The audience may doubt a consequence's desirability as well as its probability. Areni and Lutz (1988) claim that assessing a consequence's desirability is easier than assessing a consequence's probability. For instance, people may find it hard to assess the probability of a new heating system lowering heating costs while it is easy for them to assess the desirability of a cost reduction. Therefore, you would expect to encounter evidence in support of a probability claim more often than evidence in support of a desirability claim.

This expectation is supported by Schellens and De Jong's study (2000). They analyzed the occurrence of evidence in twenty Dutch Public Service Announcement documents. With respect to evidence supporting desirability they conclude that the desirability of certain consequences (e.g. self-assurance, cheerfulness) has to speak for itself and no evidence is provided (or necessary). In this paper, we focus on the evidence that can be provided to support probability claims.

1.2 Different types of evidence

Claims about the probability of a consequence can be supported by several types of evidence. This paper distinguishes between four types of evidence: anecdotal, statistical, causal, and expert evidence. Following Rieke and Sillars (1984, p. 92), we consider anecdotal evidence as the presentation of examples and illustrations. Statistical evidence is the numerical compacting of specific instances (Rieke & Sillars, 1984, p. 94). The use of causal evidence implies the prediction (or explanation) of a certain event on the basis of causal relations. Expert evidence constitutes citing an authority's opinion (Rieke & Sillars, 1984, p. 94).

The choice of these four evidence types is based on two lines of reasoning. First, the (relative) persuasiveness of each of these evidence types has already been the subject of empirical research. A review of this research is given in the next paragraph. By focusing on these same types of evidence, the predictive and explanatory force of the results is stronger.

The second line of reasoning is that the types of evidence resemble the major classes of research techniques in social sciences. Elmes, Kantowitz and Roediger (1989, p. 15) state that science is a valid way to acquire knowledge about the

world around us, i.e. scientists produce knowledge about the probability of events. As such, their research provides evidence to support probability claims about the occurrence of events.

Elmes et al. (1989, p. 17) distinguish between three major classes of research techniques: observation, correlation, and experimentation. A prime example of observation is the case study: people are interviewed in depth to acquire descriptive histories. As such, the case study provides anecdotal evidence. Correlation is the assessment of the co-occurrence of certain variables, e.g. voting behavior and socio-economic status. These correlations are often based on large-scale surveys that produce statistical evidence. Experimentation aims at explaining the relationship between variables in terms of cause-and-effect. A successful experiment provides causal evidence. Expert evidence is not produced by a research technique, but it is often used in reporting on research. It is common practice to support claims in introductions of research papers by referring to experts who have forwarded the claim.

1.3 Empirical research on the relative persuasiveness of different types of evidence

A number of studies examined the relative persuasiveness of different types of evidence. The most frequently studied comparison is the one between anecdotal and statistical evidence. Baesler and Burgoon (1994) review the results of 19 studies in which the persuasiveness of anecdotal evidence was compared to that of statistical evidence. In 13 studies, anecdotal evidence proved to be more persuasive than statistical evidence, in 2 studies the reverse pattern was found, and the remaining 4 studies reported no significant differences. This review seems to indicate quite clearly that anecdotal evidence is more persuasive than statistical evidence.

However, Baesler and Burgoon (1994) note that in many of the studies the number of cases represented is not the only difference between anecdotal and statistical evidence. The anecdotal evidence was often longer, more comprehensible, and more vivid than the statistical evidence. Baesler and Burgoon conducted an experiment in which they manipulated the number of instances that were reported on in the evidence while keeping the length, comprehensibility, and vividness of the anecdotal and statistical evidence constant. Their results showed that the statistical evidence was more persuasive than the anecdotal evidence as long as the other factors were kept constant. These results suggest that the results of studies showing that anecdotal evidence

is more persuasive than statistical evidence should be interpreted as showing that vivid evidence is more persuasive than pallid evidence.

Hoeken (2001a) tried to replicate the results reported on by Baesler and Burgoon. He had participants rate the probability of the claim that “putting extra streetlights on the sidewalks in the town of Haaksbergen would result in a sharp decrease of the number of burglaries”. This claim was supported either by statistical evidence or by two types of anecdotal evidence. The statistical evidence referred to a study conducted by the Dutch organization of municipalities on the results of putting extra streetlights on sidewalks in 48 Dutch towns. This had led to a 42% average decrease in the number of burglaries. The anecdotal evidence referred to the experiences of another Dutch town in which putting extra streetlights on the sidewalk led to a decrease in the number of burglaries by 42%. This other town resembled Haaksbergen either very much (similar anecdotal) or very little (dissimilar anecdotal evidence).

The participants in this study rated the similar and dissimilar anecdotal and statistical evidence as equally vivid. They rated the dissimilar anecdotal evidence as less relevant than the similar anecdotal evidence and the statistical evidence. The latter two did not differ from each other with respect to the relevance ratings. However, the acceptance of the probability claim that installing additional streetlights would result in a sharp decrease in the number of burglaries, was not influenced by the type of evidence provided. That is, regardless of whether the evidence was statistical, anecdotal similar or anecdotal dissimilar, the claim was accepted to the same degree.

It seems contradictory that participants were equally persuaded by the similar and the dissimilar evidence despite the fact that they considered the latter type of evidence less relevant. This discrepancy may be the result of the order in which the different questions were asked. Their opinion on the evidence’s relevance was asked after they had rated the acceptability of the claim. Possibly, the participants only realized that the dissimilar evidence was less relevant when asked to reflect upon it.

The results of Hoeken (2001a) differ from those of Baesler and Burgoon (1994). Whereas the latter report that statistical evidence is more persuasive than anecdotal evidence when controlling for vividness, the former reports no difference in persuasiveness between the two types of evidence. Perhaps this result was caused by the fact that upon a superficial inspection, all types of

evidence contain statistical information (a decrease of the number of burglaries by 42%). For the anecdotal evidence, this percentage is based upon the experience of one town, for the statistical evidence the percentage is the average percentage of 48 towns. The participants may have overlooked this difference.

In a follow-up study, Hoeken (2001b) again compared the relative persuasiveness of anecdotal and statistical evidence. In this experiment, the claim was about the probability of a multifunctional cultural center in the town of Doetinchem being profitable. Once again, the statistical evidence consisted of a study by the Dutch organization of municipalities on the profitability of such centers in a large number of Dutch towns. The anecdotal evidence consisted of the experience of the town of Groningen with the multi-functional cultural center that had been built there. Groningen is very dissimilar to Doetinchem thereby constituting dissimilar anecdotal evidence. This was done because the similarity of the evidence had no effect in the previous study.

Apart from the anecdotal and statistical evidence, Hoeken (2001b) introduced a third evidence type: causal evidence. The causal evidence consisted of three reasons why the cultural center in Doetinchem would be a success. First, many citizens from Doetinchem went to visit a cultural center far away. Second, a popular movie theatre in a nearby town had burned down. It was believed that the visitors would find their way to the cultural center in Doetinchem. Finally, Doetinchem's demographics showed that the number of well-educated, wealthy people had increased. Such people like to visit cultural centers. Hoeken hypothesized that the causal evidence would be more persuasive than either the anecdotal or the statistical evidence. He based the hypothesis on a study conducted by Slusher and Anderson (1996).

Slusher and Anderson (1996) studied the relative persuasiveness of statistical and causal evidence. The claim that Aids is not spread by casual contact was either supported by statistical or by causal evidence. The statistical evidence contained information such as "A study of more than 100 people in families where there was a person with AIDS without the knowledge of the family and in which normal family interactions such as hugging, kissing, eating together, sleeping together, etc., took place revealed not a single case of AIDS transmission." The causal evidence contained information such as "The AIDS virus is not concentrated in saliva. The virus has to be present in high concentration to infect another person and even then, it must get into that person's bloodstream." Slusher and Anderson's participants were more convinced by the causal evidence than by the

statistical evidence.

Hoeken (2001b) compares the relative persuasiveness of anecdotal, statistical, and causal evidence. The use of statistical evidence to support the claim about the cultural center's profitability proved more convincing than the use of either causal or anecdotal evidence. The latter two did not differ from each other. As in the previous study, participants were asked to rate the strength of the evidence. In this case, the anecdotal evidence was rated as weaker than the statistical and causal evidence (with the latter two not differing from each other). Again, there was a discrepancy between the perceived and actual persuasiveness. Whereas the causal evidence was just as ineffective as the anecdotal evidence in altering the acceptance of the claim, it was considered stronger than the anecdotal evidence. Again, the relative strength of the causal evidence may only have become apparent when the participants were asked to reflect upon it.

Hoeken's results (2001b) diverge from those of Slusher and Anderson (1996). Whereas the latter reported the causal evidence to be more convincing than statistical evidence, Hoeken found the reversed pattern. A closer inspection of the causal evidence used by Slusher and Anderson revealed that this evidence appears to be provided by scientists. For instance, you need special equipment and training to assess the concentration of a virus in saliva. The causal evidence used in the experiment by Hoeken may have lacked such a scientific aura. Perhaps this explains the difference in results.

1.4 The research question

The studies discussed above appear to reveal that evidence types may differ with respect to their relative persuasiveness. However, the results are rather equivocal. Which type of evidence is more persuasive than the other differs from study to study. The interpretation of the results is even more complicated because in each study only one claim and one instantiation of an evidence type were used. Consequently, a significant difference implies only that when the same claim and the same evidence are rated by a similar number of participants (who resemble the original ones), a similar difference would be obtained. A significant difference between one instantiation of different evidence types cannot be generalized to other instantiations of the same evidence type.

In this study, we have tried to extend the results of these previous studies in two ways. First, we compared directly the relative persuasiveness of four types of evidence: anecdotal, statistical, causal, and expert evidence. To our knowledge, this constitutes the first experiment in which persuasiveness of these four

evidence types is compared directly. Second, we had participants rate a large number of claims (instead of just one). Anecdotal, statistical, causal, and expert evidence was developed for each claim. Each participant rated the acceptability of claims that were supported by anecdotal evidence, by statistical evidence, by causal evidence, and by expert evidence. This research set-up enables us not only to generalize our findings over participants, but also over the types of evidence used in our study. Our research question was:

Are there differences in the persuasiveness of anecdotal, statistical, causal, and expert evidence?

2. Method

To answer the research question, twenty probability claims were presented to 160 participants. These claims were either presented without evidence, or with one of the four types of evidence. The participants rated the probability that the consequence expressed in the claim would occur.

2.1 Material

The material consisted of twenty claims in which a relation was stated between the presence of an attribute and the occurrence of a consequence, for instance:

Ontspanningsruimtes in kantoren leiden tot een sterke daling van het ziekteverzuim in Nederland.

(Relaxation rooms in offices lead to a sharp decline in absenteeism due to illness in the Netherlands.)

Four types of evidence were construed for each claim. The anecdotal evidence typically consisted of the experience of one individual, for instance:

Thomas Kepers werkt in een groot kantorenpand in de randstad. Sinds hij gebruik maakt van de gezamenlijk relax-ruimte op de tweede verdieping van zijn kantoor, heeft hij zich nooit meer ziek gemeld.

(Thomas Kepers works in a large office in the Randstad conurbation. He has not had to call in sick since he started using the relaxation room on the second floor.)

The statistical evidence consisted of a numerical summary of large number of cases, for instance:

Van 1990 tot 2000 werd er een grootschalig onderzoek gedaan naar de effecten van ontspanningsmogelijkheden op het werk. Bij bedrijven die deze voorzieningen boden, bleek 24 % minder ziekteverzuim voor te komen.

(From 1990 till 2002, a large-scale study was conducted on the effects of relaxation facilities at work. In companies that offered such facilities, absenteeism

due to illness occurred 24% less often.)

The causal evidence consisted of providing relationships that could explain why the consequence had to occur, for instance:

Door relax-kamers op het werk te plaatsen, kunnen werknemers een uur per dag ontspannen tijdens hun drukke werkzaamheden. Daardoor vermindert de werkdruk en de stress van de werknemers, wat leidt tot gezondere werknemers.

(Providing relaxation rooms at work means employees are able to relax from their demanding job for an hour a day. As a result, pressure and stress caused by work decreases in these employees, which leads to more healthy employees.)

Finally, the expert evidence consisted of quoting an authority who only rephrased the claim, for instance:

Prof. de Boer is een vooraanstaand onderzoeker op het gebied van vrije tijd en bedrijf. In zijn boek "Ruimte voor ontspanning" brengt hij naar voren dat het inrichten van ontspanningsruimten het aantal zieken vermindert.

(Professor de Boer is a leading scholar in the field of leisure time and organization. In his book Room for Relaxation he claims that the provision of relaxation rooms decreases the number of people taking time off from work due to illness.)

For each of the twenty claims, four different types of evidence were constructed along the lines described above. The different types of evidence were kept approximately equally long.

2.2 Participants

A total of 160 participants took part in the experiment. The number of women was slightly higher (84) than the number of men (76). Their ages ranged from 17 to 85, with a mean of 29.7. The level of education ranged from high school to university (master's degree).

2.3 Design

Each participant rated the probability of twenty claims, four of which were supported by anecdotal evidence, four by statistical evidence, four by causal evidence, four by expert evidence, and four claims that were not supported by any evidence. Using a latin square design ensured that each participant rated the probability of all twenty claims whereas each claim in combination with each type of evidence was rated by an equal number of participants. This resulted in five different versions of the experimental booklet. The order in which the claims were presented was identical in each of the versions.

2.4 Instrumentation

The twenty claims (and evidence) were presented in an experimental booklet. After each short text, the claim was repeated and participants had to indicate on a 5-point scale how likely they regarded the consequence to occur, for instance:

Relaxation rooms in offices lead to a sharp decline in absenteeism due to illness in the Netherlands.

Very unlikely 1 2 3 4 5 Very likely

Then, they were asked to indicate their opinion on the comprehensibility of the reasoning.

I find this reasoning

Very hard to understand 1 2 3 4 5 Very easy to understand

The latter question was only asked if the claim was supported by evidence. It was included to check whether certain types of evidence were considered more complex than others. At the end of the booklet, the participant's age, sex, and highest level of education were asked for.

2.5 Procedure

The participants were rail passengers. They were asked whether they were willing to participate in a study which would inquire about their opinion on several issues. If they agreed to participate, they randomly received an experimental booklet and any questions were answered. After they had filled out the booklet, they were told about the study's goal. On average, filling out the booklet took about 15 minutes.

Table 1 The mean probability and comprehensibility scores (1 = very hard to understand, 5 = very easy to understand) as a function of evidence type (standard deviation between parentheses).

Evidence type	Probability	Comprehensibility
Anecdotal	0.13 (0.72)	3.91 (0.90)
Statistical	0.51 (0.69)	3.98 (0.80)
Causal	0.41 (0.66)	4.02 (0.73)
Expert	0.39 (0.62)	3.93 (0.78)

Table 1 The mean probability and comprehensibility scores (1 = very hard to understand, 5 = very easy to understand) as a function of evidence

type (standard deviation between parentheses)

3. Results

The probability of claims that were not supported by evidence was rated by 32 participants. The average score of these 32 participants provided a baseline for assessing the contribution of the different types of evidence to the acceptance of the claim. To that end, the average score of the no-evidence condition was subtracted from the scores in the other conditions (evidence present). Table 1 displays the average difference scores as a function of the different evidence types together with the comprehensibility scores for these evidence types.

For the probability scores, a main effect of evidence type was found for the analysis by participants ($F_1(3, 157) = 10.77, p < .001, \eta^2 = .17$) as well as for the analysis by stimuli ($F_2(3, 17) = 12.44, p < .001, \eta^2 = .69$). Planned comparisons revealed that anecdotal evidence led to lower (difference) scores than statistical, causal, and expert evidence (analysis by participants: all p -values $< .001$; analysis by stimuli: $p < .001, p < .01, p < .05$, respectively). For the analysis by participants, there was a trend towards statistical evidence being more persuasive than causal and expert evidence, but these comparisons did not reach conventional levels of significance ($p = .09, p = .07$, respectively). For the analysis by stimuli, these trends did not arise ($p = .23, p = .16$, respectively)(**i**). There were no effects of evidence type on comprehensibility ($F_1(3, 157) = 1.68, p = .17; F_2(3, 17) = 1.03, p = .40$).

4. Discussion

The question whether there are differences with respect to the persuasiveness of the evidence types can be answered affirmatively. As was the case in the studies reported on by Baesler and Burgoon (1994) and Hoeken (2001b), anecdotal evidence proved less persuasive than statistical evidence. In contrast to the results obtained by Hoeken (2001b), causal evidence was more persuasive than anecdotal evidence in this experiment. For the first time, the persuasiveness of expert evidence was compared directly to that of anecdotal, statistical, and causal evidence. It proved to be more persuasive than anecdotal evidence, and equally persuasive as statistical and causal evidence.

An important difference with the previous studies is that the number of claims and evidence types that have been rated make it possible to draw more general conclusions with respect to the relative persuasiveness of different evidence

types. If we were to present another set of twenty claims and evidence types to this same set of participants, it is highly probable that the anecdotal evidence would once again be less persuasive compared to the other types of evidence. Noteworthy is the size of this effect ($\eta^2 = .69$). More than two-thirds of the variance in the probability ratings can be ascribed to the different types of evidence. This is a very large effect implying that argument type is a strong predictor of the acceptance of a claim.

These results appear to warrant the claim that anecdotal evidence is less persuasive compared to the other evidence types. However, this conclusion may need some modification. Depending on the type of claim, anecdotal evidence can lead to two different argumentation schemes. When anecdotal evidence is used to support a general claim about a class of events, this yields a symptomatic argumentation scheme. When anecdotal evidence is used to support a specific claim about a specific event, this yields a similarity argumentation scheme (van Eemeren & Grootendorst, 1992). The persuasiveness of anecdotal evidence may differ as a result of the argumentation scheme it is part of.

An example may clarify this. Consider the following anecdotal evidence: Thomas Kepers works in a large office in the Randstad conurbation. He has not had to call in sick since he started using the relaxation room on the second floor. In the experiment, this type of evidence was used to support general claims such as: Relaxation rooms in offices lead to a sharp decline of absence through illness in the Netherlands. Such a combination of type of evidence and type of claim yields a symptomatic argumentation scheme. The same anecdotal evidence also could have been used to support the claim that a colleague of Thomas Kepers should use the relaxation room. In that case, the anecdotal evidence is part of a similarity argumentation scheme, and the more similar the colleague and Thomas Kepers are, the more persuasive the evidence should be.

The results of the current study indicate that anecdotal evidence is, in general, less persuasive than statistical, causal, and expert evidence. The question remains whether this difference is restricted to anecdotal evidence that is part of a symptomatic argumentation scheme. In argumentation theory, people are warned against drawing general conclusions on the basis of insufficient observations. The participants in this experiment appear to have taken this warning seriously. The question whether anecdotal evidence that is part of a similarity argumentation scheme is as persuasive as the other types of evidence, remains unanswered. We hope to address this in further research.

NOTES

[i] In this study, a within-participants design was employed. Such a design may suffer from a carry over effect. Rating the probability of earlier claims and evidence, may influence the rating of subsequent claims (as participants learn to “read” the different types of evidence). To check as to whether such a carry over effect had occurred, the ratings for the first four claims based on the different evidence types were compared to the ratings of the last four claims. If the participants had learned to read the different types of evidence over the course of the experiment, their ratings of the last four types of evidence would show a larger effect than their ratings of the first four types of evidence. However, this interaction between evidence type and position in the booklet (first, last) was not significant ($F < 1$).

REFERENCES

- Baesler, J. E., & Burgoon, J. K. (1994). The temporal effects of story and statistical evidence on belief change. *Communication Research*, 21, 582-602.
- Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (Vol. 5, pp. 3-39). Hillsdale, NJ: Erlbaum.
- Eemeren, F. H. van, & Grootendorst, R. (1992). *Argumentation, communication and fallacies. A pragma-dialectical perspective*. Hillsdale, NJ: Erlbaum.
- Eemeren, F. H., van, Grootendorst, R., Snoeck-Henkemans, F., et al. (1996). *Fundamentals of argumentation theory. A handbook of historical backgrounds and contemporary developments*. Mahwah, NJ: Erlbaum.
- Elmes, D. G., Kantowitz, B. H., & Roediger III, H. L. (1989). *Research methods in psychology* (3rd. Ed.). St Paul, MN: West.
- Hoeken, H. (2001a). Convincing citizens: The role of argument quality. In D. Janssen & R. Neutelings (Eds.), *Reading and writing public documents* (pp. 147-169). Amsterdam: Benjamins.
- Hoeken, H. (2001b). Anecdotal, statistical, and causal evidence: Their perceived and actual persuasiveness. *Argumentation*, 15, 425-437.
- McGuire, W. J. (2000). Standing on the shoulders of ancients: consumer research, persuasion, and figurative language. *Journal of Consumer Research*, 27, 109-114.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion. Central and peripheral routes to attitude change*. New York: Springer.
- Rieke, R. D., & Sillars, M. O. (1984). *Argumentation and the decision making process* (2nd Ed.). New York: Harper Collins.

Slusher, M. P., & Anderson, C. A. (1996). Using causal persuasive arguments to change beliefs and teach new information: The mediating role of explanation availability and evaluation bias in the acceptance of knowledge. *Journal of Educational Psychology*, 88, 110-122.