

Quantity versus quality

Effects of argumentation in bad news letters

Daniël Janssen^{°*} and Frank Jansen[°]

[°]Department of Language and Communication, Utrecht University | The Netherlands

^{*}Department of Management, University of Antwerp | Belgium

Abstract: Do the quality and the quantity of arguments have an impact on the evaluation of bad news messages? To answer this question, two experiments were carried out using bad news letters in which the quality and the quantity of the arguments were manipulated in a contextually realistic way. The results of both experiments show that adding argumentation has a positive impact on the perceived politeness and the persuasive force of the letters. Furthermore, the studies show that the impact of strong arguments is greater than that of weak arguments. The effect of adding successive arguments is positive as well. However, the results indicate that one or two arguments are sufficient. Adding a third argument only minimally contributes to better evaluations.

Keywords: argumentation; effects; bad news; evaluation



Janssen, D., Jansen, F. (2013). Quantity versus quality. Effects of argumentation in bad news letters. *Journal of Writing Research*, 4(3), 301-331. <http://dx.doi.org/10.17239/jowr-2013.04.03.4>

Contact: Daniël Janssen, UiL-OTS, University of Utrecht, Trans 10, 3512. JK Utrecht | The Netherlands - Daniel.Janssen@let.uu.nl

Copyright: Earli | This article is published under Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported license.

1. Introduction

One of the major challenges in business writing is communicating a clear message while, at the same time, keeping the client happy and preserving the image of the organization (De Jong, 2002; Janssen 2007; Rothschild & Burnett, 1997). While clear communication may be a relatively easy task when relaying positive messages, it may become challenging when business writers have to present news that has unfavorable consequences for themselves, their organizations, and their clients. There are good reasons why bad news communication in general and bad news texts in particular have received broad attention in textbooks (e.g., Bovée & Thill, 2000; Locker, 1995), from teachers of (business) writing, as well as from writing and text researchers (cf. Brent, 1985; David & Baker, 1994; Hagge & Kostelnick, 1989; Jansen & Janssen, 2010; 2011; Janssen & Jaspers, 2004; Limaye, 1998; 2001; Locker, 1999; Schryer, 2000; Sydow Campbell, 1989; Stephen et al., 2005).

Writers of bad news messages have to perform a delicate balancing act in their attempt to reconcile the wants of the receiver with the interests of the sender, especially since the interests of senders and receivers may conflict in the case of negative messages (Jansen & Janssen, 2010). More specifically, senders face the task of presenting the bad news clearly without jeopardizing their own reputation and without diminishing the receiver's *face* more than necessary (Schryer, 2000). Researchers in the field of bad news and crisis communication are driven by the desire to understand how business writers manage to achieve this outcome, and what text features determine the effectiveness of bad news messages. In this study, we take up the main result of Jansen and Janssen (2010): of all the positive politeness strategies that were tested, only 'give reasons' (or give *arguments*) turned out to be an effective strategy within the context of bad news letters. This outcome opens up an array of interesting research opportunities. The first question that emerges has to do with the quality of the arguments. Do readers value a bad news message more highly if an unfavorable decision is accompanied by strong reasons or arguments than if the decision is accompanied by weaker ones?

The second question we would like to answer is: does the number of arguments or reasons matter? Jansen and Janssen's findings indicate that combining or 'stacking' – as it is known – different politeness strategies has no effect on readers' evaluations, since there was only an effect of a single strategy, namely 'Give reasons'. However the question remains whether combining reasons or arguments has any effect on the reception of bad news. In this study, we will fill this gap by stacking several instances of "Give reasons".

In the next section, we provide the theoretical background for the study that led to our hypotheses. We will describe two experiments with insurance claim denials as instances of bad news letters, followed by conclusions and discussion.

2. Theoretical background

There are two theoretical orientations relevant for this study: politeness theory and argumentation/persuasion theory. Politeness studies are relevant because “Give reasons” is one of the positive politeness strategies that Brown and Levinson (1984) identify. The use of arguments or reasons indicates that the receiver recognizes the sender is trying to establish a positive relationship with him or her; this is essential in (business) communication. Persuasion and argumentation theories are relevant because the addition of reasons or explanation in a bad news message is meant to explicate the rationality behind the decision. The added reasons are offered to convince the receiver that the sender has made an understandable — although negative — decision.

In terms of politeness theory, negative decisions such as insurance claim denials and job refusals are “face threatening acts” (FTAs) (Brown & Levinson, 1987). After all, when a client files a claim, he does so in good faith (or at least with some hope) that the insurance company will comply by replying that his claim is eligible for payment. Unfortunately, in the case of a denial, the insurance company does the opposite. The company writes a letter informing the client that his request is denied. In Brown and Levinson’s terms, this message harms the reader’s face. Brown and Levinson suggest that this situation can be improved upon by adding politeness strategies to the refusal. They offer a typology of politeness strategies, one of their categories being *positive* politeness strategies in which the sender conveys to the receiver that they are *cooperators*. “Give reasons” (or explanations or argumentation) is one of those positive politeness strategies.

In the context of business writing, several authors have observed that professional writers must comply with the demands of the communicative situation and, at the same time, preserve the client’s positive and negative face while maintaining their own (Pilegaard, 1997; Hagge & Kostelnick, 1989; Janssen & Jaspers, 2004; Ding, 2006). Also, several studies have demonstrated that professional writers use politeness strategies to achieve their communicative goals intentionally (Hagge & Kostelnick, 1989; Van der Mast & Janssen, 2001). In their analysis of accountants’ letters, Hagge & Kostelnick (1989) observe that accountants use several forms of indirectness (negative politeness strategies) and vagueness (an off-record strategy) to point out the ‘problems’ they have detected during an audit (= bad news):

Customer’s safekeeping records are not always updated for current charges. We noted some receipts in the customer safekeeping register for items that had been returned to the customers without record of the return (Hagge & Kostelnick, 1989, p. 324).

Van der Mast (1999) and Van der Mast and Janssen (2001) observed many instances of politeness when they studied policy writers *at work*. Parts of the documents that had to be revised to meet objections by earlier critics contained more *polyphonic text characteristics* than the originals. Although Van der Mast and Janssen and Van der Mast do not use politeness theory as their point of departure, their examples can be easily

analyzed in terms of politeness strategies. Policy writers consciously use shifts in perspective, vagueness, impersonalization, and other rhetorical devices to make their texts work (cf. Barghiela-Chiappini & Harris, 1996).

Other researchers have found instances of positive politeness strategies in corpora of professionally written bad news texts. Jablin and Krone (1984) discovered positive politeness strategies in rejection letters to job applicants. Pilegaard (1997) found positive politeness strategies (alone or in combination) in letters in which business partners negotiate an order, i.e. in letters where one might expect refusals to occur frequently. Manno (1999) gives many examples of those strategies in his corpus of French-Swiss job refusals. And Kok (1993) found combinations of two or more positive strategies in all of the letters in a corpus of 31 Dutch job refusals. In all but one of her letters “Give reasons” was one of the strategies.

Van Waes and Van Wijk (2000) studied the effects of positive politeness strategies (and a combination of negative strategies as another independent variable) on the evaluation of product recalls, another bad news genre. In one of their experiments, they found a positive effect of positive politeness strategies on the evaluation of the text, on the image of the issuing company, and on the *acceptance* of the bad news. The only study to measure the effects of “Give reasons” in combination with other positive politeness strategies is the aforementioned study by Jansen and Janssen (2010) who conducted a series of experiments with bad news letters (e.g., claim denials and job refusals) in which they manipulated the number and the type of politeness strategies. They measured the effects of their manipulation on relationship variables (such as sender’s attributes, perceived empathy, organizational image) and persuasion variables (e.g., compliance). Jansen and Janssen concluded that the combined results of the two experiments offer a clear picture: “Give reasons” has a *positive* effect on the evaluation of the letters, while two other (positive politeness) strategies have *no effect* at all (p. 2531). These outcomes are relevant for a number of reasons. First, they introduce questions about the validity of advice given in many textbooks: namely, to use positive politeness strategies. Second, they present a typology of possible differential effects of the successive addition of strategies. Jansen and Janssen distinguish an “additive model” (each strategy added has a positive effect on the reader’s evaluation), a “saturation model” (the positive effect of each successive strategy is declining), and an optimum model (the addition of one [or two] strategies initially enhances the effect, but addition of extra strategies diminishes the effect). Since the positive politeness strategies deployed in Jansen and Janssen’s experiment turned out to be ineffective in use, it was not possible to interpret the results as evidence for or against one of those models.

Third, the results suggest questions concerning the validity of politeness theory; in particular, these are more about the cognitive status of politeness theory since the presence of strategies does not seem to affect readers’ responses. Jansen and Janssen tentatively explain the outcomes of the experiments with the help of Petty and Cacioppo’s (1984) *Elaboration Likelihood Model*. The central idea in this model is that persuasion takes place in an “elaboration continuum that ranges from low elaboration

(shallow processing of the information) to high elaboration (deep processing of the information)". At the extremes of this continuum are the *central* and the *peripheral* route of persuasion. Central route processes involve careful scrutiny of the argumentation for the proposition/ conclusion. Under these conditions, readers' evaluate whether the argument itself is probable (grounded on facts), whether the argument can contribute to the proposition and whether there is no other evidence bearing on the proposition that has been overlooked. In the peripheral route, readers' employ other strategies to evaluate the proposition in the message. They do not elaborate on the argumentation as much but rely more on peripheral cues to determine their standpoint. Those peripheral cues may consist of illustrations, the reputation of the sender or the number of arguments mentioned in the text. In the latter case readers' do not scrutinize the argumentation but count the number or arguments and evaluate this aspect by a rule of thumb: the more the better (of a good thing, that is) (cf. Petty & Cacioppo, 1984; Chaiken, 1987; Marshall et al., 2002).

The results of Jansen and Janssen (2010) indicate that their participants took the central route. This would explain why only "Give reasons" (= argumentation) caused the effect and other strategies did not. In this study, we will pursue this path by incorporating argument quality in the experiment. If participants evaluate the text with strong arguments more favorably than the text with weaker argumentation, we know for sure that they, indeed, processed the text (at least the argumentative parts of it) along the central route.

Of course, we are not the first researchers to be interested in argument quality. Reinard (1988) presents an impressive overview of research into the effects of evidence in persuasive communication. Reinard's reviews, for instance, study the effects of different *types* of evidence; e.g., statistical versus examples and opinions, factual versus testimonial. He concludes, "aside from the previously described research by Petty and his associates indicating that empirical research findings and statistical information are especially persuasive among subjects who are personally involved with the topics, differences created by various evidence forms have been elusive" (p. 25).

In a more recent study on the effects of anecdotal, statistical, and causal evidence, Hoeken (2001) found text with statistical evidence to be most persuasive and text with anecdotal evidence to be as unconvincing as text with causal evidence. However, it seems relevant for the evaluation of this result that the participants had indicated beforehand that they considered anecdotal evidence, as such, to be less convincing. Van Eemeren, Garssen, and Meuffels (2009) found evidence that language users consider fallacies less acceptable than non-fallacious argumentation, but they did not measure the actual effect on the persuasiveness of messages. Finally, in her comprehensive study, Šorm (2010) showed that normatively strong arguments are not necessarily more persuasive than normatively weak arguments. For one thing, the effects depend on the type of argumentation; i.e., Šorm finds strong argumentation from cause to effect to be more effective than weak argumentation of the same type.

However, it turns out that the same does not hold true for argumentation from authority (p. 230).

In most of these studies, however, the effect of argumentation is measured in a realistic context. Participants in the experiments were presented separate argumentations (argument + conclusions) that were evaluated in terms of argument quality. Although this approach has given us many valuable insights in the effects of argumentation quality, it is still unclear whether we would find the same results would we present argumentation in realistic documents, such as business letters. The research regarding “evidence” (statistical, anecdotal) is not directly transferable to our work because the argumentation in the bad news letters is not purely *evidence-based*, but also rule-based as we shall see in the next sections.

For our study, we are interested in the way argumentation affects the acceptance of bad news. However, as most research in this field points out that strong arguments are often more effective than weak arguments, it seemed reasonable to expect arguments in bad news messages to work out the same way.

With respect to the number of arguments, the study of Petty and Cacioppo (1984) showed that participants who had little interest in the topics were more influenced by the number of arguments than were people with high involvement with the topic. Other studies mentioned in Reinard (1988) were more or less inconclusive: “on the whole, research on the quantity of arguments indicates that no magic number exists for evidence use” (p. 40).

Finally, we could not find any research that studied the interaction between the number of arguments and the quality of argumentation. As there is a reasonable chance that a single strong argument functions as a *knock-out* factor and brings about a ceiling-effect, we expected weak arguments to benefit more from the presence of other weak arguments than strong arguments to benefit from additional strong arguments.

In this study, we will first attempt to replicate the finding of Jansen and Janssen by answering the following questions:

1. Are bad news letters with argumentation as effective as bad news letters without argumentation?

The leading hypothesis is:

- 1a. Bad news letters with argumentation are more effective than bad news letters without argumentation.

Second, we will try to explain the effects in terms of processing by answering the following two questions:

2. Are texts with strong arguments as effective as texts with weak arguments?

The hypothesis is:

- 2a. Texts with strong arguments are more effective than texts with weak arguments.

3. Are texts with one argument as effective as texts with more than one argument?

The leading hypothesis is:

- 3a. More arguments make a text more effective.

As far as the combination of quality and quantity is concerned, we have no clear hypothesis. We will approach the issue of the interaction between quality and quantity of argumentation in an explorative manner.

3. Experiment 1: Quantity and quality of argumentation

In the first experiment we explored the effects of the number of arguments and the quality of arguments on the evaluation of bad news.

3.1 Method

As in Jansen and Janssen (2010; 2011), we used insurance claim denials as typical examples of bad news letters. The main reason for this choice was that our previous experiments had demonstrated this form of bad news communication generated results that were both (ecologically) valid and of practical use for teachers and practitioners. Furthermore, refusal letters are ubiquitous and all participants in this study had at least some experience with reading them. We assumed that the participants' experience would make it easier for them to empathize with the receiver of the refusal letter. Also, this type of letter is usually short (one page), which makes it suitable for stimulus material. Finally, using the same type of stimuli enabled us to evaluate the outcomes of this experiment in the broader context of earlier research. We will come back to this point when we discuss the results.

In our experiment, we based the letters on real examples from Kok (1993) and from a major insurance company in the Netherlands. All texts were formatted in the conventional style of business letters, with logo and address information. The mean length of the letters was 186 words (min. 180, max. 191).

The names of the insurance company sending the letters were fictitious so that a company's established image would not affect readers' evaluations (see appendix A for an example of the body one of the four Dutch letters, which has been translated into English.)

Independent variables

As mentioned earlier, we wanted to replicate the effects found in Jansen and Janssen (2010) and measure a difference in evaluation between letters with and letters without reasons or argumentation ("Bald On-record" in terms of Brown & Levinson, 1987).

Furthermore, we needed to make a distinction between weak and strong arguments. Although the empirical focus of this study prevented us from elaborating on theoretical problems in normative argumentation theory a few words are in order about the way

we selected weak and strong arguments. As far as we know two ways to determine the quality of argumentation have been used in experiments like this study. The first way is normative argumentation theory (Van Eemeren & Grootendorst, 1992), which distinguishes between reasonable argumentation, following the rules for rational discussions, and reasonable use of argumentation schemes and the so-called fallacies (Van Eemeren, Garssen, & Meuffels, 2009). This theory seemed less apt for our purposes for reasons of ecological validity: insurance letters do not explain a reimbursement denial with the help of traditional fallacies (cf. Woods & Walton, 1982). Furthermore, the distinction between, for instance, a fallacious use of an argumentation based on authority and a reasonable use of the same argumentation scheme is not always easy to make on strictly normative grounds. This distinction is especially so when readers need to take specific contexts into consideration. Within the context of claim denials all arguments are essentially identical from a normative perspective: namely, variants of the rule-based argumentation scheme (see Schellens & De Jong 2004). Following Toulmin's (1985) terminology (see also, Lunsford 2002) this scheme is:

Data / argument: The client's claim has characteristic X (for example, it is more than one year old)

Warrant / major premise: Policy terms state that claims with characteristic X are denied

Claim / conclusion: Therefore: we have to deny your claim

As this scheme holds for all explanations in reimbursement denial letters, it is not possible to define quality in terms differences in normative argumentation schemes unequivocally.

The second way to distinguish strong from weak arguments is empirically. This method is applied successfully in the ELM experimental tradition where argument quality is based on the opinions of the target group itself. In this tradition, the selection of the arguments takes place in two phases. In the first phase researchers make an inventory of possible weak and strong arguments. In the second phase they present the arguments to a sample of the same population they plan to use for their experiment, with the request to select the most appealing and also the most unconvincing arguments. The selected arguments are subsequently used in the experiment. This approach was less fruitful for our purpose because the maximal differentiation of arguments with respect to their quality leads in the case of the weak arguments to arguments that are so weak that it is doubtful that they will be considered as arguments at all.

Therefore, we used a slightly different approach. Although we used the second phase when we pretested the argumentation in our experiment, we preselected the strong and weak arguments by applying Schellens (1987) norms for rule-based

argumentation (see also Schellens and De Jong, 2004) in combination with a third criterion that we called *normative-analytical*.

We considered arguments *strong* if three conditions were met:

1. The warrant (the rule on which the decision was based) was made explicit.
2. The explicit warrant referred to the terms of the insurance policy directly, that is, without addition of hedges expressing the views or interpretations of the insurance company.
3. The receiver could apply the rule strictly and unambiguously, that is, the argument for the denial was directly related to the policy terms and there was no question about the application of the rule in the specific case (cf. norms for rule-based argumentation in Schellens & De Jong, 2004).

In the case of weak arguments, one or more of the three conditions above were not met. To find valid and realistic cases of argumentation in the field of insurance, we asked 15 major insurance companies to send us form letters and examples of claim denials — more specifically, denial of travel insurance claims. From the responses, we selected 11 arguments and divided them into strong and weak arguments based on the aforementioned criteria. The three examples of strong argument in the case of a travel insurance claim were:

1. Your claim is more than one year old. (Policy terms state that claims have to be filed within a year of the alleged theft.)
2. You did not file a police report. (Policy terms state that claims can only be submitted after a police report has been filed.)
3. Theft took place outside Europe. (Your travel insurance policy only covers damage within the E.U.)

In the following example, we illustrate the relevant policy terms, showing the steps from facts (data) to claim denial:

Data / argument: Your claim is more than one year old.

Warrant / major premise: Policy terms state that a claim has to be submitted within a year of the alleged theft.

Claim / conclusion: Therefore, we have to deny your claim.

The weak arguments that were used in the letters were:

1. You have claimed a higher amount than the actual value of the laptop computer. According to our information, the price of the laptop was € 678,30.
2. The form provides evidence that you left your laptop unattended near the swimming pool. Therefore, you are no longer entitled to reimbursement.
3. Our records indicate that you have already filed three claims this year.

The first thing to notice about the weak arguments is that – although they are weaker than the strong arguments – they are not fallacious in the traditional meaning of the word ‘fallacy’. Nevertheless, they can be considered weaker because in none of these instances the relationship between the argument and the claim denial is made *explicit* with a warrant quoting the relevant policy term (the normative analytical criterion). Furthermore the weak arguments give rise to valid counter-argumentation, which is the basic criterion for weaker argumentation in Schellens (1987) and Garsen (2001) approaches to evaluating argumentation. In argument (2) and (3) the writer indicates that the grounds on which the decision was based left openings for interpretation and debate (‘the form *provides evidence*’ and ‘our record *indicate*’). Argument (1) is debatable because prices of consumer goods fluctuate and differ from shop to shop and if the claim is considered too high, the insurance company can reimburse a lower amount instead of simply denying the claim all together. Finally, argument (2) is weaker because policy terms (and thus warrants in the argumentation) are never this specific. Insurance policies state, for instance, that ‘owners should take good care of all items under insurance’. In such instances it is debatable whether a dive in the pool and leaving your laptop on the table falls within the scope of this term.

We are aware of the fact that this normative-analytical approach to argument quality contains elements of subjectivity. To test whether this analytical point of view could be substantiated by empirical data, we did a small pretest among 15 participants aged 24-50. The results showed that our strong arguments were indeed valued more than the weaker arguments.

The second independent variable was the number of arguments. We designed six letters in total in which we combined either the strong or the weak arguments. In other words, there were no conditions with combinations of weak and strong arguments. So, in the experiment, the strong condition consisted of claim denial letters with one, two, or three strong arguments, and the weak condition consisted of letters with one, two, or three weak arguments. The most elaborate version with three strong arguments was as follows:

After careful examination of your claim, we regrettably have to inform you that we cannot reimburse the damage for three reasons. The first reason is that your claim is more than one year old. Policy terms state that you lose coverage if you do not file a claim within a year of the alleged theft. The second reason is that you did not file a report of the alleged theft at the local police station as is required in our policy terms and conditions. The third reason is that your insurance policy does not cover damage outside the E.U.

Based on this information we have decided not to reimburse you.

The version with no arguments contained only the decision; the version with one argument contained the decision and the first-mentioned reason; and the version with two arguments included the decision and the first two reasons. To underline the differences between the versions with one, two, or three arguments, we marked the

enumeration using argumentation indicators (cf. Van Eemeren et al., 1989) in two ways. First, we forewarned the reader with “for X reason(s)” (cf. Kamalski et al., 2008). Second, we used “the first reason”, “the second reason”, “the third reason” since this is generally considered the most marked signal for a list structure (cf. Sanders & Noordman 2000). Of course, we followed the same path in the weak arguments condition. The most elaborate version with three weak arguments was the following:

After careful examination of your claim we regrettably have to inform you that we cannot reimburse the damage for three reasons. The first reason is that – in our opinion – you have claimed more than the damage; our information shows that the price for that specific laptop computer was € 678,30. The second reason is that the form provides evidence that you left your laptop computer unattended near the swimming pool; therefore, you are no longer entitled to reimbursement. The third reason is that our records indicate that you have filed three previous claims this year.

Based on this information, we have decided not to reimburse you.

Again, the version with one weak argument only contained the first reason; the version with two arguments included the first two reasons, etc.

We maximized the chance that participants evaluating the conditions with two or more arguments (both weak and strong) interpreted the argumentation as independent in convergent argumentation, which is also known as multiple argumentation in normative argumentation theory (Snoeck Henkemans, 2000). We tried to achieve this interpretation by introducing each argument by the standard indicator for this type of argumentation, namely *the (first/second/third) argument is ...* (Snoeck Henkemans 2000). Because the argumentative context of insurance claim denial makes the interpretation of arguments as linked (or complementary coordinative) almost impossible, as each breach of the policy condition provides sufficient reason for imbursement denial. We cannot exclude, however, that readers may have interpreted the arguments as convergent (cumulative). This is an empirical question that we will address in the discussion.

All letters had a direct structure: first the decision, and then the arguments for it. The internal order of the arguments was fixed as well.

Dependent variables

For this study, we used an evaluation instrument that had proved to be reliable in earlier experiments (Jansen & Janssen, 2010). Again, we refrained from presenting any direct questions about strength and quantity of arguments in the questionnaire (available via WritingPro.eu). We believed that it is essential for participants to be ignorant of the researchers’ goals — not only to prevent socially correct answers, but also to prevent unwanted framing on the part of the participants. We did not want them to focus on argumentation or lack thereof. Instead, we asked our subjects to assess the overall quality of the letters. Furthermore, we were interested in their assessment of the expressive and relational aspects of the letters because they form strong indicators for

the persuasive strength of the message. Finally, we measured the perceived persuasive strength more directly with a proposition regarding the expected compliance and the perceived reasoning.

For instrumentation, we used a questionnaire by which the participants could evaluate the letter (see Appendix A). First, the participants gave a general assessment of the overall quality of the message: a report mark (ranging from 1 to 10 as is the convention in the Netherlands) for the entire message. Then the participant answered two open questions measuring their initial response to the message. For the more specific evaluation, we used seven-point Likert scales (ranging from 1 totally disagree to 7 totally agree).

Some of the propositions were formulated positively, while other propositions were formulated negatively as is the custom in this type of research to avoid the danger of semi-automatic responses. Before the participants filled in the scales, they gave a general assessment of the overall quality of the letter: a report mark (ranging from 1 to 10 as is the convention in the Netherlands) for the entire text. The Likert scales were followed by three questions about the participants' demographics: in particular, gender, education, and age. To check for memory effects, the questionnaire closed with a question about the number of reasons or arguments in the letter, a question about the participants' experience with insurance matters ("did you ever file an insurance claim?"), and a question about the ecological validity of the letter ("this letter may have been sent to a customer").

The Likert scales were initially designed to measure the effects of our manipulation on perceived politeness (friendliness and empathy) and persuasiveness (compliance and persuasiveness) of the claim denial. To test this assumption we performed a factor analysis (principal component with Varimax rotation) that resulted in four components explaining 64% of the variance in total. However, Cronbach's alpha of the two items in the fourth component was only .51 so we had to neglect this cluster in our further analyses. The other three components were composed from items with factor loads $>.50$.

The three remaining components match two politeness variables and a persuasion variable. The initial distinction we made between persuasiveness and compliance was not reflected in the data.

To measure the effects on persuasion we used the following scale:

1. Persuasiveness, e.g., *The arguments of the writer are convincing; The reader would probably comply with the decision* (Cronbach's $\alpha = .81$)

To measure the effects on perceived politeness we used two scales:

2. Friendliness, e.g., *The writer is helpful; The writer is kind* (Cronbach's $\alpha = .72$)
3. Empathy, e.g., *The writer is interested in me; The writer shows involvement* (Cronbach's $\alpha = .79$)

Participants

Of the 175 participants in this experiment, 88 were men (50 %) and 87 were women (50%). We thus had 25 participants in each condition. The age range was from 23 to 73, with an average age of 38 (SD 14,7). Randomization checks showed that participants were distributed equally in respect to their gender and age across all the conditions. Of the participants, 83 (47%) had finished some form of secondary vocational education; 61 (35%) had obtained a bachelor's degree, and 30 (17%) a master's degree. A One-way ANOVA revealed no effect of education on any of the dependent variables ($p > 0.05$). Most participants (61.7%) had hands-on experience with insurance claims. All participants had a job or have had a job, thus they had, at least, an elementary knowledge of insurance matters.

They all had a Dutch cultural background and were all competent native speakers and readers of Dutch. No participant was paid for his or her contribution. Most participants volunteered for the experiments during a train ride as they commuted to work; others were recruited from within family circles.

Design and procedure

The experiment had a between-participants 2x4 design for both quality and quantity variables (weak versus strong arguments, zero, one, two, or three arguments, and weak versus strong arguments). That is, each participant evaluated a letter in one condition. Twenty five participants evaluated each letter.

First the participants read the scenario:

In March 2006 you took a trip to Sri Lanka. Before you left you purchased travel insurance at Solar Travel Insurances Ltd.

During your two-week holiday, on March 14, 2006, your Acer Aspire 9301 AWSMi laptop computer was stolen near the swimming pool of your hotel.

More than a year later, on June 15, 2007, you fill in a form on which you report the stolen laptop computer. On the form you mention the date and location of the loss and a value of € 1475.69.

On June 20, 2007 you receive the following letter from Solar Travel Insurance Ltd.

Of course the letters varied with the condition (see: *Independent variables*). However, the opening and closing paragraphs were identical in all conditions.

Opening

Dear Sir or Madam,

On June 15th we received your form on which you report theft of your Acer Aspire 9301 AWSMi laptop computer. In your letter you mentioned that you lost your laptop on March 14, 2006, during a trip to Sri Lanka, near the swimming pool. You also mentioned a value of € 1475,69.

After the receipt of your letter we examined the data in light of the terms and conditions of your insurance policy to see if your case is eligible for reimbursement.

Body

[no reasons, one strong/weak reason, two strong/weak reasons, three strong/weak reasons]

Closing

We regret having to tell you this. If you disagree with this decision, you can lodge a written complaint against this decision. Please send your complaint before July 30, 2007 to Solar Travel Insurance, Customer Service, Postbox 16520, 2500 KB The Hague. For more information, please check the enclosed brochure "You do not agree with a decision?" that is included in the envelope with this letter or available at www.solar.nl.

We hope we have informed you satisfactorily.

Yours sincerely,

(...)

All letters had an authentic layout that followed the conventions of business letters. They were about 170 words long (see Appendix B). The participants were given instructions stating that they had just claimed the missing item. Then they read the claim denial completely unaware of the specific content of the questionnaire. Reading the letter took on average five minutes. Once they had finished reading, the participants filled out the questionnaire. If they felt it was necessary, they could reread the letter before making any assessments. On average, the participants needed 13 minutes to complete the task.

3.2 Results

What are the effects of argument quality and quantity on the evaluation of the bad news letters? Table 1 presents the mean scores and standard deviations of the report mark and the other dependent variables for presence versus absence of argumentation.

Table 1. Means scores and standard deviations (*SDs*) of the clusters of dependent variables (1 = negative evaluation, 7/10 = positive evaluation), according to argument presence

	- Arguments	+ Arguments
Overall (report mark)	5.6 (1.4)	7.0 (1.1)
Politeness variables		
Friendliness	3.8 (1.3)	4.6 (1.1)
Empathy	2.8 (1.1)	3.6 (1.2)
Persuasion variable		
Persuasiveness/compliance	2.3 (1.3)	4.8 (1.1)

The results clearly show a main effect of argumentation presence. The average scores of the letter without argumentation are lower than the average scores on the claim denials with argumentation. A multivariate analysis of variance revealed significant effects of the presence of argumentation. Using Hotelling's trace statistics, there was a significant effect of argumentation on the dependent variables ($T = .46$, $F_{(4, 169)} = 19.3$, $p < .001$, $\eta^2 = .31$). Separate ANOVA's showed effects on overall evaluation ($F_{(1, 172)} = 30.9$, $p < .001$, $\eta^2 = .15$), friendliness ($F_{(1, 173)} = 11.2$, $p = .001$, $\eta^2 = .06$), empathy ($F_{(1, 173)} = 11.4$, $p = .001$, $\eta^2 = .06$) and persuasiveness ($F_{(1, 173)} = 64.3$, $p < .001$, $\eta^2 = .27$).

Thus, adding argumentation to bad news led to a more positive evaluation in terms of politeness and persuasion. Now that we know that argumentation matters, it becomes more interesting to see whether the quality of the argumentation had any effect. To determine such an effect, we split the second column and performed another MANOVA (see Table 2 for the results).

Table 2. Means scores and standard deviations (SDs) of the clusters of dependent variables (1 = negative evaluation, 7/10 = positive evaluation), according to argument quality

	- Arguments	Weak	Strong
Overall (report mark)	5.6 (1.4)	6.8 (1.0)	7.2 (1.1)
Politeness variables			
Friendliness	3.8 (1.3)	4.4 (1.3)	4.8 (0.9)
Empathy	2.8 (1.1)	3.4 (1.2)	3.9 (1.1)
Persuasion variable			
Persuasiveness/compliance	2.3 (1.3)	4.4 (1.5)	5.1 (1.2)

In our analyses we used argument strength as our factor and the overall evaluation, politeness and persuasion as dependent variables. We report the mean scores on the bald-on-record letter only as a point of reference. We did not include the claim denials without argumentation in this analysis as this analysis would partially replicate the effects found in Table 1.

The results shown in Table 2 indicate that the quality of argumentation did make a difference. The scores on all dependent variables were higher in the strong argument condition than in the weak argument condition. A multivariate analysis of variance confirmed an effect of argument strength. Using Hotelling's trace statistics, we measured a significant effect of argumentation strength on the dependent variables ($T = .102$, $F_{(4, 144)} = 20.1$, $p < .01$, $\eta^2 = .09$). Separate ANOVA's revealed that argument quality affected the overall evaluation ($F_{(1, 147)} = 4.1$, $p = .04$, $\eta^2 = .03$), the perceived friendliness ($F_{(1, 147)} = 5.9$, $p = .02$, $\eta^2 = .04$) and empathy ($F_{(1, 147)} = 8.3$, $p = .005$, $\eta^2 = .05$), as well as the persuasiveness ($F_{(1, 147)} = 10.8$, $p = .001$, $\eta^2 = .07$), of the bad news letter.

When we added the letters without argumentation in our multivariate analyses and performed an additional Post Hoc Tukey HSD we found that the difference between the

mean scores of the letter with strong arguments differed significantly from the mean scores on the bald versions on all dependent variables ($p < .001$). The difference between the mean scores of the version without argumentation and version with weak argumentation was only significant on persuasion ($p < .001$) and the overall evaluation ($p < .001$).

Now that we know more about the effects of the quality of arguments, we can focus on the quantity. Table 3 presents the mean scores and standard deviation for the overall evaluation and the clusters of dependent variables for the number of arguments.

Table 3. Means scores and standard deviations (*SDs*) of the clusters of dependent variables (1 = negative evaluation, 7/10 = positive evaluation), according to number of arguments

	- Arguments	One	Two	Three
Overall	5.6 (1.4)	7.1 (1.0)	6.9 (1.1)	6.9 (1.3)
Politeness variables				
Friendliness	3.8 (1.3)	4.6 (0.9)	4.5 (1.3)	4.7 (1.0)
Empathy	2.8 (1.1)	3.4 (1.1)	3.6 (1.1)	3.9 (1.2)
Persuasion variable				
Persuasiveness/compliance	2.3 (1.3)	3.9 (1.5)	4.9 (1.1)	5.5 (1.1)

Again, we added the means for the bald-on-record letter for reference purposes only. We analyzed the data excluding letters without argumentation for reasons mentioned earlier. Again, multivariate analyses of variance revealed a significant effect. Using Hotelling's trace statistics, we measured a significant effect of argumentation quantity on the dependent variables ($T = .367$, $F_{(8, 284)} = 6.5$, $p < .001$, $\eta^2 = .16$). Separate ANOVA's revealed only one significant effect, on persuasiveness ($F_{(3, 171)} = 21.0$, $p < .001$, $\eta^2 = .21$). No other significant effects were found. A post hoc Tukey analysis pointed out that letters with two and three arguments were considered more persuasive than letters with a single argument. The difference between two and three arguments was not significant. In other words, adding more than two arguments did not make the letter more convincing.

Furthermore we found an interaction effect of argument quality x quantity on one of the politeness variables, Friendliness ($F_{(2, 144)} = 4.6$, $p < .01$, $\eta^2 = .06$). Adding a second weak argument had a significantly negative impact on the perceived friendliness of the claim denial ($M_{\text{one}} = 4.5$; $SD = 0.8$), $M_{\text{two}} = 3.8$; $SD = 1.5$). Adding a third weak argument led to the same evaluation as adding a single argument ($M_{\text{three}} = 4.7$; $SD = 1.3$). In the condition with strong arguments this effect did not occur. The number of arguments had no effect on the perceived politeness.

3.3 Conclusion

In this experiment, one of the effects from the study by Jansen and Janssen (2010) was replicated. The outcomes indicate that adding reasons or argumentation to validate bad news had a strong positive effect on readers' evaluations. Letters with argumentation (strong and weak) were perceived as more polite and more persuasive. The presence of argumentation also led to higher general appreciation of the text.

The results also indicate that the quality of the arguments in bad news letters matters as well. When we compared the effects strong versus weak arguments in the claim denial, the strong arguments had more impact in persuading the reader to accept the decision. The image of the organization benefitted from the presence of strong arguments, as well. Organizations are perceived as friendlier and more empathic when they add strong arguments instead of weak arguments to their bad news. However, in this study these effects were smaller than the effects of argumentation presence in general. Effect sizes for presence argumentation were considerably larger than those for argumentation quality.

When we investigated the effects of the number of arguments, we observed a positive general effect of adding one argument to a bald-on-record claim denial. The positive effect of adding a second argument was more limited. Generally, the number of arguments affected only the persuasiveness of the letter and not the overall evaluation or the perceived politeness. Adding a second argument made the letter more persuasive, adding a third argument was of no additional value. Finally, the data showed that adding a second weak argument was detrimental to perceived politeness. The outcomes of this first experiment strongly indicate that our participants processed the message in terms of the elaboration likelihood model more centrally than peripherally. We found a systematic difference in the evaluation of strong and weak arguments and no systematic differences between one, two and three arguments, which are both indications of central processing according to Petty and Caccioppo (1984). Nevertheless we needed to do a second argument for two reasons. As the experiment used only one message we could not rule out that the effects found were limited to this specific letter. Therefore, we conducted a second experiment with the same design, using a different bad news letter. Furthermore, we wanted to gain more insight in the amount of elaboration and needed more data that provided us more direct information on the characteristics of participants' reading processes.

4. Second experiment: Replication and processing

The second experiment served two purposes. For reasons of external validity we wanted to know whether the findings from the first experiment could be replicated using a different letter. Second, we wanted to gather more information on the reading processes of readers of bad news messages. More specifically, we needed more information on how they processed argumentation.

4.1 Method

The dependent variables as well as the procedure were almost identical to the first experiment. As far as the material (independent variable) is concerned, we made a number of changes. We used identical argumentation and the same genre of bad news letters. However, we changed the subject from a stolen laptop computer to an Olympus digital camera that represented a lower value (€ 149), and we altered the holiday destination from Sri Lanka to Egypt. The sender was changed to Global Travel Insurances Ltd, and we adapted the letterhead accordingly. The letters were of comparable layout and identical length to the letters used in the previous experiment. Again, we opted for a between-participants 2x4 design (weak and strong arguments, and the presence of zero, one, two, or three arguments).

A second alteration in the design was that the participants in this second experiment received an additional task beyond reading the instruction and evaluating the letter. We used a method called *thought listing* to gather more information about the *route* our participants followed when processing the information in the letter: central or peripheral (Petty & Cacioppo, 1986). Thought listing is a method first developed by Brock (1967) and Greenwald (1968) and successfully applied and slightly amended by Cacioppo and Petty (1981) and Cacioppo, Von Hippel, and Ernst (1996). Thought listing is an *open question* method to measure and categorize a participant's personal evaluation of a text. Cacioppo, Von Hippel, and Ernst (1996) refer to these evaluations as "mental contents": all measurable verbalizations of individual thoughts, feelings, ideas, expectations, judgments, and representations. Petty and Cacioppo's research shows that participants will display more cognitive responses in situations with high involvement, indicating more elaboration. The central idea behind the thought listing method is simple: the more relevant ideas participants remember and verbalize, the deeper the processing during reading must have been. To avoid anticipation, the participants were given the instruction for thought listing after reading the instruction and the letter, but before filling out the questionnaire. They were asked to put the letter and the questionnaire aside. Then they received a form with five empty boxes and the following instruction:

Fill the boxes on the form with all thoughts that come to mind with respect to the letter you just read. Just write your first thought in the first box, the second thought in the second box, etcetera. There are no right or wrong answers.

Do not feel obliged to fill in all the boxes, and don't worry about spelling or grammar.

You will have about five minutes.

The forms were scored by a research assistant in collaboration with one of the authors who was unaware of the conditions the participants were assigned to. The participants' remarks on the forms were coded according to valence (are they positive = "good letter", negative = "very accusatory", neutral = "I would reread the terms and conditions"), content (do remarks concern the content of the letter; e.g., "the arguments

are irrelevant," or other aspects; e.g., "very direct"), and agreement (are the remarks in agreement with decision in the letter; e.g., "my claim was denied legitimately," or not; e.g., "why don't they reimburse € 155 then!"). The coders worked to a point of full consensus about the scoring.

The questionnaire in this experiment was identical to the one used in the previous study. The reliabilities were - again - high: *friendliness* (Cronbach's $\alpha = .80$), *empathy* (Cronbach's $\alpha = .79$), *persuasiveness* (Cronbach's $\alpha = .87$).

Participants

In total 94 men (41.8%) and 131 women (58.2%) participated in the experiment.¹ Ages ranged from 20 to 75 with an average age of 35.4 (*SD* 13.7). Randomization checks revealed that participants were distributed equally with respect to their gender and age across all the conditions. Of the participants, 26.3% had finished some form of secondary vocational education, 33.7% had obtained a bachelor's degree, and 40% a master's degree. Three participants did not provide information about their level of education. Again, all levels of education were equally spread over the conditions. Most participants (64%) had hands-on experience with insurance claims. All participants had a job or have had a job, and they had experience with travel insurance. Of all participants, 200 made a thought list, 100 in the strong and 100 in the weak condition. They all had a Dutch cultural background and were all fluent native speakers and readers of Dutch. No participant was paid for participating. Most participants volunteered for the experiments during a train ride between two major cities as they commuted to work; others were recruited from within family circles by one of our research assistants.

4.2 Results

When we examined the results of the second experiment, we noticed small differences and apparent similarities between the two studies. In our presentation of the data we follow the same organization as in Experiment 1. The effects of the presence of argumentation are presented in Table 4.

The data show the same pattern as in the first experiment. Again the mean scores of the letter without argumentation were lower than the mean scores on the claim denials with argumentation. Multivariate analyses of variance revealed significant effects of the presence of argumentation. Using Hotelling's trace statistics, there was a significant effect of argumentation on the dependent variables ($T = .303$, $F_{(4,152)} = 11.5$, $p < .001$, $\eta^2 = .23$). However separate analyses of variance revealed significant effects of the presence of argumentation only on overall evaluation ($F_{(1,155)} = 12.3$, $p = .001$, $\eta^2 = .02$) and persuasiveness ($F_{(1,155)} = 37.8$, $p < .001$, $\eta^2 = .20$). The effect of argumentation presence on empathy almost reached significance ($F_{(1,155)} = 3.4$, $p = .06$, $\eta^2 = .02$), but the effect on friendliness was nowhere near significance ($F_{(1,155)} = 2.5$, $p = .12$, $\eta^2 = .02$).

Table 4. Means scores and standard deviations (*SDs*) of the clusters of dependent variables (1 = negative evaluation, 7 = positive evaluation), according to argument presence

	- Arguments	+ Arguments
Overall (report mark)	6.0 (1.2)	7.1 (1.3)
Politeness variables		
Friendliness	4.0 (1.3)	4.7 (1.3)
Empathy	2.8 (1.0)	3.3 (1.2)
Persuasion variable		
Persuasiveness/compliance	2.4 (0.9)	4.7 (1.8)

Thus, claim denials with argumentation are valued more highly in terms of overall evaluation and persuasion, but they are not considered more 'polite' in the sense of Brown and Levinson. What about the effects of the quality of argumentation? To gain insight in the effect of quality we split up the second column and performed another MANOVA (see Table 5 for the results).

Table 5. Means scores and standard deviations (*SDs*) of the clusters of dependent variables (1 = negative evaluation, 7/10 = positive evaluation), according to argument strength

	- Arguments	Weak	Strong
Overall (report mark)	6.0 (1.2)	6.7 (1.3)	7.5 (1.1)
Politeness variables			
Friendliness	4.0 (1.3)	4.3 (1.2)	5.1 (1.3)
Empathy	2.8 (1.0)	2.9 (1.1)	3.5 (1.4)
Persuasion variable			
Persuasiveness/compliance	2.4 (0.9)	3.6 (1.7)	5.7 (1.2)

In our analyses we used argument strength as our factor and the overall evaluation, politeness and persuasion as dependent variables. We report the mean scores on the bald-on-record letter only for reference purposes. We did not include the claim denials without argumentation in this analysis because this would (partially) replicate the effects found in Table 1.

The results in Table 5 indicate that the quality of argumentation did make a difference. The scores on all dependent variables were higher in the strong argument condition than in the weak argument condition. A multivariate analysis of variance confirmed an effect of argument strength. Using Hotelling's trace statistics, we measured a significant effect of argumentation strength on the dependent variables ($T = .713$, $F_{(6, 338)} = 20.1$, $p < .001$, $\eta^2 = .26$). Separate analyses of variance confirmed that

argument quality affects the overall evaluation ($F_{(1, 127)} = 17.5, p < .001, \eta^2 = .12$), the perceived friendliness ($F_{(1, 127)} = 37.4, p < .001, \eta^2 = .23$) and empathy ($F_{(1, 127)} = 5.0, p = .03, \eta^2 = .04$), as well as the persuasiveness ($F_{(1, 127)} = 64, p < .001, \eta^2 = .11$) of the bad news letter. When we compared the mean scores in the weak and strong condition to the bald-on-record condition using a MANOVA and additional Post Hoc Tukey HSD, we found that the letters with strong arguments differed significantly from the letters without arguments ($p < .001$), whereas the weak arguments had no effect on the dependent variables.

Now that we know more about the effects of the quality of arguments, we can focus on the quantity. Table 6 presents the mean scores and standard deviation for the overall evaluation and the clusters of dependent variables for the number of arguments. When we examined the effects of the number of arguments, we found results similar to our first experiment (see Table 6).

Table 6. Means scores and standard deviations (*SDs*) of the clusters of dependent variables (1 = negative evaluation, 7/10 = positive evaluation), according to argument quantity

	- Arguments	One	Two	Three
Overall (report mark)	6.0 (1.2)	6.9 (1.2)	7.3 (1.0)	6.9 (1.6)
Politeness variables				
Friendliness	4.0 (1.3)	4.8 (1.2)	4.7 (1.3)	4.5 (1.4)
Empathy	2.8 (1.0)	3.3 (1.0)	3.4 (1.3)	3.2 (1.3)
Persuasion variables				
Persuasiveness/compliance	2.4 (0.9)	3.9 (1.9)	4.8 (1.7)	5.0 (1.6)

Again we added the means for the bald-on-record letter for reference purposes only. Again multivariate analysis of variance revealed a significant effect. Using Hotelling's trace statistics, we measured a significant effect of argumentation quantity on the dependent variables ($T = .442, F_{(9, 503)} = 20.1, p < .001, \eta^2 = .128$). A separate analysis of variance only revealed one significant effect, that is, on persuasiveness ($F_{(2, 127)} = 7.9, p = .001, \eta^2 = .11$). No other significant effects were found. A post hoc Tukey analysis showed that letters with two and three arguments were considered more persuasive than letters with a single argument. The difference between two and three arguments was not significant. In other words adding more than two arguments did not make the letter more convincing and two arguments was more convincing than one. We found the same pattern in Experiment 1. However, in this second experiment we did not find any interaction effect between the number and the quality of arguments.

4.3 Conclusion

In this second experiment, many of the effects found in the first experiment were replicated using a different letter. Again the outcomes indicate that adding reasons or

argumentation to validate bad news has a positive effect on readers' evaluations. The effect in this second experiment was, however, more limited to persuasion and general evaluation. We could not replicate the effects on politeness variables in Experiment 1 and earlier studies by Jansen and Janssen (2010).

The results on the quality of the arguments in bad news letters matters were identical to the outcomes of the first experiment. When we compared the effects strong versus weak arguments in this claim denial, the strong arguments had more impact than weak arguments on persuasion, overall evaluation and perceived politeness. Organizations are perceived as friendlier and more empathic when they add strong arguments instead of weak arguments to their bad news letters.

The effect of the number of arguments was – again - limited. The number of arguments in general only affected the persuasiveness of the letter and did not influence the overall evaluation or the perceived politeness. Adding a second argument made the letter more persuasive, adding a third argument was of no additional value.

4.4 Thought listing results

The purpose of this second experiment was – besides replicating the first – to gather more information about the way readers process argumentation in bad news letters. The results of the thought-listing procedure helped us to gain more insight into how the different text versions were processed.

Table 7. Thought listing results according to presence of argumentation

	- Arguments	+ Arguments
<i>Content/non content</i>		
Remarks about content	0 (0%)	125 (63%)
No remarks about content	25 (100%)	75 (27%)
<i>Agreement with decision</i>		
Agree	2 (8%)	81 (40%)
Disagree	2 (8%)	73 (37%)
Neutral	21 (84%)	46 (23%)
<i>Valence</i>		
Positive	13 (14%)	147 (21%)
Negative	37 (40%)	158 (23%)
Neutral	43 (46%)	392 (56%)
<i>Total</i>	93	697

Table 7 also shows presence versus absence of argumentation as a factor for the content scores. Here we see that the presence of argumentation gives rise to significantly more remarks ($Chi^2 = 13.28$; $df = 2$, $p = 0.01$) after reading the letter. Furthermore the proportions in Table 7 indicate that adding argumentation urged the readers to take sides for or against the insurance company's decision.

Did the presence of argumentation influenced the readers' focus? The results in Table 7 indicate that the participants who evaluated letters with argumentation were significantly more focused on the content of the text than were the readers of the Bald On-record letter (without argumentation) ($Chi^2 = 32.67$; $df = 1$, $p < 0.01$), leading to more distinct opinions. The readers of letters without argumentation were often more neutral towards the conclusion, where readers of the letters with argumentation expressed more articulate standpoints against or in favor of the decision in the text ($Chi^2 = 39.55$; $df = 2$, $p < 0.01$).

When we split the + argumentation category into weak and strong argumentation, we get an even more specific view of the effects of argumentation on the processing of the text. When we compare the frequencies in Table 8 with the ones that could be expected under H0 we see that strong arguments gave rise to 15 percent more positive and 34 percent less negative remarks on the thought-listing forms. The situation for the weak argument condition was exactly the opposite, namely 16 percent less positive and 37 percent more negative comments. All differences are significant as $Chi^2 = 26.42$; $df = 2$, $p < 0.001$.

Table 8. Thought listing results according to strength of argumentation

	Strong	Weak
Content/non content		
Remarks about content	69 (69%)	56 (56%)
No remarks about content	31 (31%)	44 (44%)
Agreement with decision		
Agree	68 (68%)	13 (13%)
Disagree	11 (11%)	62 (62%)
Neutral	21 (21%)	25 (25%)
Valence		
Positive	88 (24%)	59 (18%)
Negative	54 (15%)	104 (31%)
Neutral	220 (61%)	172 (51%)
<i>Total</i>	362	335

Another question concerned the relationship between argument quality and the focus of the readers. It is obvious from Table 8 that the remarks were equally distributed over both conditions, so we were unable to measure any significant differences with respect to the aspects of the letters that the thought-listing remarks referred to, content or non-content. The third question was about the influence of argument quality on agreement (see again Table 8).

The data in Table 8 reveal a strong effect of the quality of argumentation on agreement. Readers of the strong arguments version made 68 percent more remarks expressing agreement with the decision, and 69 percent fewer remarks indicating they disagreed. In the weak argument condition, the picture was exactly reversed. The results are also very significant ($Chi^2 = 73.32; df = 2, p < 0.01$).

The thought-listing data thus showed that the presence of argumentation and the quality of argumentation had an effect on the number and kind of remarks that the participants made. These results indicate that the argumentation in the texts is indeed processed deeply, and that differences between the conditions should be attributed to differences in elaboration.

5. General conclusions and discussion

The results suggest that presenting arguments for a decision that implies bad news for the receiver is generally a good idea. Both experiments show that bad news with argumentation is valued more than bad news without explanation. The results of the two experiments are in line with hypothesis 1a, "bad news letters with argumentation are more effective than bad news letters without argumentation," and give a satisfactory replication of Jansen and Janssen's (2010) finding that adding argumentation is an effective strategy. In Experiment 1 we measured effects on the overall evaluation, the perceived politeness and the persuasiveness of the letter, indicating that argumentation not only led to more compliance but also to a better client-organization relationship. In the second experiment, this effect on politeness could not be replicated. However, we did find effects of strong argumentation on perceived politeness. This absence of a more generic politeness effect may be explained by a difference in imposition. In the second experiment the rejected claim was € 145, while in the first experiment the rejected claim was € 1475. When the imposition is higher one would expect more effect of politeness strategies, such as "give reasons." Further experiments should determine if this explanation holds.

Finally, thought-listing data from the second experiment revealed that adding argumentation led to more elaboration and thus deeper processing.

Turning our attention to argumentation quality, the results indicated that readers value letters with strong arguments more highly than letters with weak arguments. In both experiments we found significant effects on the overall evaluation and the persuasiveness, as well as on the politeness. In both experiments, strong arguments were evaluated more positively than weak arguments. Effects were measured on the

overall evaluation, the persuasiveness and the perceived politeness of the letter. Again, these outcomes were further substantiated by the result in the thought-listing data. The strong arguments led to significantly more positive remarks and signs of agreement and compliance. These results are consistent with hypothesis 2a, "texts with good arguments are more effective than texts with weak arguments."

Argument quantity had a demonstrably positive effect on readers' evaluations as well, although this effect was limited to persuasion. Readers found denial letters with two arguments more convincing than letters with a single argument. Adding a third argument did not have any effect. Adding more than one argument did not affect the overall evaluation of the letters or the perception of relationship between the insurance company and its client.

However, our research had a number of limitations. First, we only used one type of bad news in one context. Therefore, the variety was limited, which makes it valuable to do future experiments on other types of bad news (e.g., turning down an invitation or a marriage proposal, refusing to lend a person a certain amount of money). A more systematic manipulation of the gravity of the bad news may be especially rewarding.

Second, experiments are inherently artificial. It remains uncertain whether our participants have reacted in the same way as readers in real life who are more involved because the bad news directly affects their personal interests. In other words, all experiments give rise to questions about the ecological validity. At the same time, we find it hard to come up with other means to gather data systematically for these kinds of research questions. Third, the design of the experiment, especially the manipulation of the argument quantity conditions (0, 1, 2 and 3 arguments), makes it difficult to ascribe the effects we found in the respective conditions to argument quantity with absolute certainty because the arguments were added in a fixed order. Therefore, we cannot rule out the alternative explanation that the differential results point to significant quality differences *within* the categories of strong and weak arguments (see also Jansen & Janssen 2010). Fifth, the specific linguistic and cultural background of our participants who are speakers of Dutch in the Netherlands may threaten the external validity of the results. Are the results equally relevant for English and international correspondence? We are inclined to answer this question affirmatively. The differences between the Dutch and the Anglo-Saxon culture are rather small. Furthermore, there is not much evidence that other cultures react differently to argumentation (but see Siegel (1991) for a different view).

6. Discussion

When we review our results, we cannot but consider them as strong indications that the participants of the experiments – receivers of bad news letters – have evaluated the letters by way of central processing rather than peripheral processing. We will mention only two of the strongest indicators. First, the consistent superior effect of strong arguments compared to weak arguments points in the direction of central processing. If

readers only considered the mere presence of argumentation to determine their evaluation, we would not have found an effect of argumentation quality. Second, the absence of consistent effects from adding a third argument argues against peripheral processing. If readers would have used a general rule of thumb such as *more is better*, we would have found an overall effect of argumentation quantity – which we did not. Furthermore consider it very likely that real receivers of bad news letters, such as claim denials, would be even more willing to process the text deeply as reading the text is in their direct interest. In our view, it is thus plausible that the variables we manipulated, affect real readers in the same way as in our experiment. Our study shows that adding (solid) argumentation to claim denials may improve - or at least conserve - the relationship between the organization and its clients and may lead to more compliance.

The second question that remains has to do with the fact that for a number of dependent variables, two arguments suffice to achieve saturation and additional arguments have no further effect. How do readers process these additional arguments? In the ELM-model, central processing of information is distinguished from peripheral. The idea is that under specific conditions information is processed via one of these routes. However, in the case of multiple strong arguments, it is not unthinkable that readers process only one of the strong arguments and decide that this single argument is sufficient to validate the conclusion. Consequently, they may feel no further need to centrally process any additional argumentation. This would explain why only one strong argument suffices to achieve saturation. Further experiments with thought listing are needed to test this hypothesis.

Finally, although our normative-analytic approach to the selection of arguments of high and low quality was successful (after all, we did find strong effects of argument quality), it has to be more thoroughly implemented in future experiments by keeping the presence of warrants constant over all conditions while varying the amount of explicitness and elaboration.

Notes

- 1 Fifty participants read letters that had an *indirect* structure: first arguments, then conclusion. In order to compare the results of this experiment with the results of the first experiment, these participants were left out of the statistical analysis (cf. Jansen and Jansen 2011).

Acknowledgements

The authors would like to thank: Mark Torrance, Jos Hornikx and two anonymous reviewers for their valuable comments; Frans van Eemeren and Hans Hoeken for their inspiration. Nhung Nguen and Sybren Visser helped us gathering data for the experiments.

References

- Bargiela-Chiappini, F., & Harris, S.J (1996). Requests and Status in Business Correspondence. *Journal of Pragmatics* 28, 635-662. doi: 10.1016/0378-2166(96)89191-0
- Bovée, C., & Thill, J.V. (2000). *Business communication today*. Englewood (New Jersey): Prentice Hall.
- Brent, D. (1985). Indirect structure and reader response. *Journal of Business Communication*, 22 (2), 5-8. doi: 10.1177/002194368502200201
- Brock, T. C. (1967) Communication Discrepancy and Intent to Persuade as Determinants of Counterargument Production. *Journal of Experimental Social Psychology*, 3, 296–309. doi: 10.1016/0022-1031(67)90031-5
- Brown, P., & Levinson, S. C. (1987). *Universals in language use: Politeness phenomena*. Cambridge (UK): Cambridge University Press.
- Cacioppo, J.T., von Hippel, W., & Ernst, J.M. (1997). Mapping Cognitive Structures and Processes through Verbal Content: The Thought-Listing Technique. *Journal of Consulting and Clinical Psychology*. 65(6), 928–940. doi: 10.1037/0022-006X.65.6.928
- Cacioppo, J. T., & Petty, R. E. (1981). Social psychological procedures for cognitive response assessment: The thought-listing technique. In T. V. Merluzzi, C. R., Glass, & M. Genest (Eds.), *Cognitive Assessment*, p. 309–342. New York: Guilford Press.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology* 39, 5, p. 752–766. doi: 10.1037/0022-3514.39.5.752
- David, C., & Baker, M. A. (1994). Rereading bad news: Compliance-gaining features in management memos. *Journal of Business Communication*, 31, 267–290. doi: 10.1177/002194369403100403
- De Jong, J. (Ed.) (2002). *Zakelijk corresponderen* [Business Correspondence]. Groningen, Netherlands: Wolters Noordhoff.
- Ding, D. D. (2006). An indirect style in business communication. *Journal of Business and Technical Communication*, 20, 87–100. doi: 10.1177/1050651905281051
- Garssen, B. (2001). 'Argument Schemes'. In F. H. van Eemeren (ed.), *Crucial Concepts in Argumentation Theory*, Amsterdam University Press, Amsterdam, p. 81–99.
- Greenwald, A. G. (1968). Cognitive Learning, Cognitive Response to Persuasion, and Attitude Change. In: A. G. Greenwald, T. C. Brock, and T. M. Ostrom (Eds) *Psychological Foundations of Attitudes*. New York: Academic Press. doi: 10.1016/B978-1-4832-3071-9.50012-X
- Hagge, J., & Kostelnick, C. (1989). Linguistic Politeness in Professional Prose: A Discourse Analysis of Auditors' Suggestion Letters, with Implications for Business Communication Pedagogy. *Written Communication* 6, 312–339. doi: 10.1177/0741088389006003004
- Hoeken, H. (2001). Anecdotal, statistical, and causal evidence: their perceived and actual persuasiveness. *Argumentation* 15, p. 425–437. doi: 10.1023/A:1012075630523
- Jablin, F.M. & Krone, K. (1984). Characteristics of Rejection Letters and Their Effects on Job Applicants. *Written Communication* 1(4), 387–406. doi: 10.1177/0741088384001004001
- Jansen, F., & Janssen, D. (2010). Effects of positive politeness strategies in business letters. *Journal of Pragmatics*, 42, 2531–2548. doi: 10.1016/j.pragma.2010.02.013
- Jansen, F., & Janssen, D. (2011). Explanations first; A case for presenting explanations before the decision in Dutch bad news messages. *Journal of Technical and Business Communication*, 25, 36-67.
- Janssen, D. (Ed.) (2007). *Zakelijke communicatie* [Translation: Business communication]. Groningen, Netherlands: Wolters Noordhoff.
- Janssen, D., & Jaspers, J. (2004). Client-Oriented Correspondence: How Do you Actually Do It? *Document Design* 12, 3, .187–189.
- Kamalski, J., Lentz, L., Sanders, T., & Zwaan, R. A. (2008). The forewarning effect of coherence markers in persuasive discourse: evidence from persuasion and processing. *Discourse Processes*, 45, 545–579. doi: 10.1080/01638530802069983

- Kok, A. (1993). *Zo'n onaangename standaardbrief. Een onderzoek naar afwijzingsbrieven aan sollicitanten* [Translation: Such an unpleasant form letter. An investigation of rejection letters to job applicants]. Unpublished master thesis Utrecht University, Netherlands.
- Limaye, M.R. (1988). Buffers in bad news messages and recipient perceptions. *Management Communication Quarterly*, 2, 90–101. doi: 10.1177/0893318988002001006
- Limaye, M.R. (2001). Some Reflections on Explanations in Negative Messages. *Journal of Business and Technical Communication*, 15, 100–110. doi: 10.1177/105065190101500105
- Locker, K. O. (1995). *Business and administrative communication*. Chicago: Irwin.
- Locker, K. O. (1999). Factors in reader responses to negative Letters: Experimental evidence for changing what we teach. *Journal of Business and Technical Communication*, 13, 5–48. doi: 10.1177/105065199901300101
- Lunsford, K.J. (2002). Contextualizing Toulmin's model in the writing classroom: A case study. *Written Communication*, 19, p. 109–174. doi : 10.1177/074108830201900105
- Manno, G. (1999). Savoir Refuser à l'Écrit: Analyse d'un Enchaînement Non Préfère de Macroactes de Discours. *Journal of French Language Studies* 9(1), 39–68. doi: 10.1017/S0959269500004440
- Marshall, R., Chuan, N.E., & Woonbong, N. (2002). Children's use of consumer heuristics. In: S.M. Broniarczyk & K. Nakamoto (Eds.), *Advances in Consumer Research*, Volume 29, p. 108–113. Valdosta GA: Association for Consumer Research.
- Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses on argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology* 46(1), p. 69–81. doi: 10.1037/0022-3514.46.1.69
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag. doi: 10.1007/978-1-4612-4964-1
- Pilegaard, M. (1997). Politeness in business requests. *Journal of Pragmatics*, 28, 223–243. doi: 10.1016/S0378-2166(96)00084-7
- Reinard, J.C. (1988). The empirical study of the persuasive effects of evidence: The status after fifty years of research. *Human Communication Research* 15(1), 3–59. doi: 10.1111/j.1468-2958.1988.tb00170.x
- Rothschild, E.H., & Burnett, R. (1997). *Business Communication*. Prentice Hall, New Jersey.
- Sanders, T. J. M., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes*, 29, 37–60. doi: 10.1207/S15326950dp2901_3
- Schellens, P. J. (1987). 'Types of Argument and the Critical Reader', in F. H. van Eemeren, R. Grootendorst, J. A. Blair and Ch. A. Willard (eds.), *Argumentation: Analysis and Practices*, p. 34–41. Foris, Dordrecht.
- Schellens, P.J., & De Jong M. (2004). Argumentation schemes in persuasive brochures. *Argumentation* 18, 295–323. doi: 10.1023/B:ARGU.0000046707.68172.35
- Schryer, C. F. (2000). Walking a fine line. Writing negative letters in an insurance company. *Journal of Business and Technical Communication*, 14, 445–497. doi: 10.1177/105065190001400402
- Siegel, H. (1999). Argument quality and cultural difference. *Argumentation* 13, 183–201. doi: 10.1023/A:1026466310894
- Snoeck Henkemans, A.F. (2000). State-of-the-art: the structure of argumentation. *Argumentation* 14, 447–473. doi: 10.1023/A:1007800305762
- 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*. Doctoral dissertation. Radboud University. Utrecht (The Netherlands): LOT.
- Stephen, K. K., Callish, P., & Bailey, C. (2005). Communicating with Stakeholders During a Crisis: Evaluating Message Strategies. *Journal of Business Communication*, 42, 390–419. doi: 10.1177/0021943605279057

- Sydow Campbell, K. (1990). Explanations in Negative Messages: More Insights from Speech Act Theory. *Journal of Business Communication* 27, 357–375 doi: 10.1177/002194369002700403
- Toulmin, S. E. (1958). *The Uses of Argument*. Cambridge MA: Cambridge University Press.
- Van der Mast, N. (1999). *Woordenwisseling; een onderzoek naar de manier waarop schrijvers consensus over beleidsteksten bewerkstelligen*. (Translation: (Ex)changes of Words; An Inquiry into the Way Writers Create Consensus on Policy Documents). Amsterdam: Thesis.
- Van der Mast, N., & Janssen, D. (2001). Strategic Revision of Public Documents: How Policy Writers Revise Documents in Order to Reach Consensus. In: D. Janssen & R. Neutelings (eds.). *Reading and Writing Public Documents*, p. 171–210. Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/ddcs.1.10mas
- Van Eemeren, F.H., & Grootendorst, R. (1992). *Argumentation, communication, and fallacies. A pragma-dialectical perspective*. Hillsdale (NJ): Erlbaum.
- Van Eemeren, F., Grootendorst, R., & Meuffels, B. (1989). The skill of identifying arguments. *Argumentation; Journal of the American Forensic Association*, 25, 239–245.
- Van Eemeren, F., Grootendorst, R., & Snoeck Henkemans, F. (1996). *Fundamentals of argumentation theory. A handbook of historical backgrounds and contemporary developments*. Mahwah: Erlbaum.
- Van Eemeren, F.H. Garssen, B., & Meuffels, B. (2009). *Fallacies and the Judgments of Reasonableness: Empirical Research Concerning the Pragma-Dialectical Discussion Rules*. Dordrecht: Springer. doi: 10.1007/978-90-481-2614-9
- Van Waes, L., & Van Wijk, C. (2000). The influence of politeness on the perception of product recall notices. *Document Design*, 2, 272–279. doi: 10.1075/dd.2.3.07van
- Woods, J., & Walton, D. (1982). *Argument: The Logic of the Fallacies*. Toronto: McGraw-Hill.

Appendix A. Scales (and factor loads)

Overall (report mark)

Rate the letter on a scale from 1-10

Politeness variables

Friendliness

The writer comes across as friendly (.58)

The writer is helpful (.81)

The writer is arrogant (recoded) (.72)

Empathy

The writer shows involvement (.61)

The writer is interested in me (.73)

The writer takes me seriously (.82)

The writer empathizes with the situation of the reader (.56)

Persuasion variable

Persuasiveness/compliance

The reader will probably comply with the decision (.89)

The decision to deny the reimbursement is valid (.69)

The arguments of the writer are convincing (.74)

The writer gives sufficient arguments (.82)

The bad news is presented convincingly (.80)

REMARK

The questionnaire itself (Dutch) is available via WritingPro.eu: [download](#)

Appendix B: Example of a letter (translated from Dutch)

Dear Sir or Madam,

On June 15th we received your form on which you report theft of your Acer Aspire 9301 AWSMi laptop computer. In your letter you mentioned that you lost your laptop on March 14, 2006, during a trip to Sri Lanka, near the swimming pool. You also mentioned a value of € 1475.69.

Upon receipt of your letter, we examined the data in light of the terms and conditions of your insurance policy to see if your case is eligible for reimbursement.

After careful examination of your claim, we regrettably have to inform you that we cannot reimburse the damage for three reasons. The first reason is that your claim is more than one year old. Policy terms state that you lose coverage if you do not file a claim within a year of the alleged theft. The second reason is that you did not file a report of the alleged theft at the local police station as is required in our policy terms and conditions. The third reason is that your insurance policy does not cover damage outside the E.U.

Based on this information, we have decided not to reimburse you.

We regret having to tell you this. If you disagree with this decision, you can lodge a written complaint. Please send your complaint before July 30, 2007 to Solar Travel Insurance, Customer Service, Postbox 16520, 2500 KB, The Hague. For more information please check www.solar.nl or the brochure "*You do not agree with a decision?*".

We hope we have informed you satisfactorily.

Yours sincerely,

(signature)