INDIANA UNIVERSITY
PURDUE UNIVERSITY
INDIANAPOLIS

Department of Economics
Working Paper Series

Persuasion in Fundraising Letters: An Interdisciplinary Study

Elizabeth Goering
Ulla M. Connor
Ed Nagelhout
Richard Steinberg

Working Paper No. 2008-04

425 University Blvd, CA 516
Indianapolis, IN 46202

Persuasion in Fundraising Letters: An Interdisciplinary Study

Elizabeth Goering
Department of Communication Studies
Indiana University Purdue University Indianapolis
425 University Blvd. CA 307F, Indianapolis, IN 46202
Phone: (317) 278-3136; Fax: (317) 278-1025
Email: bgoering@iupui.edu

Elizabeth Goering is an Associate Professor in the Department of Communication Studies at IUPUI. The primary focus of her research is the intersection between communication and culture in a variety of contexts, including organizational, health, and media.

Ulla M. Connor
Professor of English
Barbara E. and Karl R. Zimmer Chair in Intercultural Communication
Director, Indiana Center for Intercultural Communication
Union Building 407, 620 Union Drive, IUPUI
Indianapolis IN 46202-5170
Phone: 317-278-2441
Email: uconnor@iupui.edu

Ulla Connor is Professor of Linguistics and Director of the Indiana Center for Intercultural Communication at Indiana University Purdue University Indianapolis. Her research focuses on the analysis of spoken and written discourses.

Ed Nagelhout
University of Nevada, Las Vegas
Department of English
4504 Maryland Parkway, Box 5011
Las Vegas, NV 89154
Phone: 702-895-5073
Email: ed.nagelhout@unlv.edu

Ed Nagelhout, a specialist in rhetoric and writing, is the Director of Business and Technical Writing at the University of Nevada, Las Vegas. His current research interests include the rhetoric of poverty and writing program administration.

Richard Steinberg
Department of Economics
Indiana University Purdue University Indianapolis
425 University Blvd. CA 307F, Indianapolis, IN 46202

Phone: (317) 278-7221
Email: rsteinbe@iupui.edu

Rich Steinberg is Professor of Economics, Philanthropic Studies, and Public Affairs at IUPUI. His research covers determinants of giving and volunteering and the behavior, effectiveness, and regulation of nonprofit organizations. He was Co-President of ARNOVA 1992-94.

Research was supported by a grant from the Center on Philanthropy at Indiana University. The authors wish to thank Alicia Anino, Bradley Andersen, Molly Anthony, Adolph Kotlik, Kyle McIntosh, and Toula Vasilopolous for research assistance and Jane Chu, Robert Sandy, Adrian Sargeant, Jen Shang, Kathy Steinberg, and participants at ARNOVA's annual conference Indiana University's Workshop in Multidisciplinary Philanthropic Studies for helpful suggestions on earlier drafts.

Persuasion in Fundraising Letters: An Interdisciplinary Study

Abstract

In this paper, we report experimental evidence on the effectiveness of several techniques of persuasion commonly utilized in direct-mail solicitation. The study is built on theory-based, descriptive models of fundraising discourse and on comparisons of recommended and actual practices related to three dimensions of persuasion: rhetorical, visual, and linguistic. The specific rhetorical variable included is persuasive appeal (rational, credibility, or affective). The visual variable selected for the study is the presence or absence of bulleted lists, and the linguistic variable included is readability, or the complexity of exposition. Subjects were presented with pairs of fictive direct-mail appeals from imaginary universities that differ in these dimensions and asked to allocate a hypothetical $100 across each pair. Results suggest that letters utilizing credibility appeals and letters written at a high level of readability produce the highest donations.

Fundraising appeals are designed to persuade individuals to look beyond narrow self-interest towards the greater social good. In this paper, we report experimental evidence on the effectiveness of several techniques of persuasion commonly utilized in direct-mail solicitation. Subjects read pairs of direct-mail appeals differing in rhetorical, visual, and linguistic dimensions, and then allocated hypothetical donations totaling $100 across each pair.

In order to select our experimental treatments, we first reviewed the advice contained in popular fundraising textbooks. Most such advice is comes from the experience-based learning of practitioners, rather than controlled scientific studies. We therefore selected specific practices for experimental testing based on theory-based, descriptive models of fundraising discourse (such as those by Upton, 2002, and Connor & Gladkov, 2004) and divergences between recommended and actual practices in the field. The rhetorical dimension included three treatments -- rational appeals (arguments based on logical reasoning through the presentation of facts and causal relationships), credibility appeals (arguments based on the authority of the writer) and affective appeals (arguments targeting the reader's emotions). The visual dimension included two treatments differing in the presence or absence of bulleted lists, and the linguistic dimension included two treatments differing in readability, the complexity of exposition. All the pairs of direct-mail solicitations were fictive, asking for donations to fictional universities. We found that letters utilizing credibility appeals and those written at a high level of readability produced the highest donations.

The next section reviews the relevant literature. Then we describe our procedure for selecting research hypotheses, our methodology, and empirical results.

Literature Review

Fundraising is largely a persuasive activity that seeks to convince donors to contribute to a worthy cause. Our focus here is on one tactic commonly used as part of an organization's fundraising campaign, the direct-mail letter. Most nonprofit organizations use direct-mail in some way, and usage has been growing exponentially in recent years (Torre and Bedixen, 1988). Warwick (2000, p. 166) argues that this fundraising method is "the single biggest means used by nonprofits to recruit new donors," noting that "research repeatedly confirms that the majority of first time gifts to charity are made by mail." Abelen, Redeker, and Thompson (1993, p. 325) concur, explaining that direct-mail letters are the "most important instrument for communicating the 'good cause' of a non-profit organization to a wide range of prospective donors."

The remainder of this section describes general research on the effectiveness of direct mail fundraising. Then we describe literature specific to each of the three dimensions of persuasion tested by our study.

What Works?

Because direct-mail letters are such an integral part of many organizations' fundraising strategy, practitioners and researchers alike share an interest in identifying factors that contribute to the success of direct-mail appeals. And indeed, researchers from a variety of disciplinary perspectives, including marketing (e.g., Basu, Basu & Batra, 1995; Berger & Smith, 1997, Diamond & Gooding-Williams, 2002), communication studies (e.g., Abrahams and Bell, 1994), economics (e.g. Andreoni and Petri, 2004), and psychology (e.g., Petty, Cocioppo & Schumann, 1983) have focused attention on this research question.[1]

Organizations often conduct their own research, test-marketing various letters to determine what works for their campaigns. However, the data, research methods, and research

findings are proprietary and, if they are shared at all, not subject to the review processes of refereed journals. But some situational analyses (where subjects make hypothetical choices after being presented with hypothetical appeals), laboratory experiments, and field experiments have appeared in the academic literature. Much of this research compared the efficacy of the presence or absence of particular fundraising strategies, such as including a suggested donation amount (Weyant & Smith, 1987; Fraser, Hite & Sauer, 1988; Desmet and Feinberg, 2003), personalizing persuasive appeals by multiple use of the potential donor's name in the solicitation (Turner & Yeakel, 1994), or including a list of other donors and the size of their contributions as a strategy for motivating compliance (Reingen, 1982). Other researchers (i.e., Stone, 1992; Tversky & Kahneman, 1981) have explored the impact on compliance of the ways in which fundraising messages are framed. For example, Tversky & Kahneman (1981) compared the impact of positive and negative frame valence, operationalized as describing outcomes in terms of "lives saved" versus "lives lost," and concluded that positively framed messages yield greater outcomes.

Warren and Walker (1991) reported on a field experiment conducted in Perth, Australia. They mailed 2648 letters of eight types, differing in three dimensions – inclusion of statements designed to induce empathy ("picture yourself in the position of the other" vs. "picture a person"), designed to create the perception that the problem was extensive (describing a family's plight vs. that of a community), and designed to create the perception that the problem was persistent (describing immediate vs. long-term needs of the victims). The confirmed their expectation that donations were more likely to be made and were larger when the problem was portrayed as less extensive and of shorter duration, but found no statistically significant effect of empathy induction. They acknowledge that the failure to find an effect of empathy induction

may be an artifact of the low response rate and consequent lack of statistical power (they were prospecting for new donors, rather than seeking new donations from established donors), highlighting the a difficulty in conducting field studies with this type of subject pool. Interestingly, they first conducted an experiment testing their procedures on 72 undergraduate Psychology majors. Results for this group, using intent to donate as the dependent variable, were quite similar to those from the field experiment.

Smith and Berger's (1996) field study used a factorial design with three factors: use of anchors, framing, and reference information. Their third factor is similar to our rhetorical factor – whether the letter contains factual/statistical information or narrative/experiential information. They found that reference information influenced the size of the gift but not the response rate. Similarly, Parsons' (2001) field study found that providing repeat donors with financial efficiency information influenced the likelihood of giving, whereas this information had no effect on prospective donors who had not donated previously. Finally, Katzev (1995) found that inclusion of vivid information about what a specific dollar contribution "buys" increased the amount but not the probability of giving. However the sample size was small and the conclusion was sensitive to inclusion of an outlier.

Other field experiments tested direct-mail packaging and enclosures. Falk (2004) found that when gifts to the prospective donor were included in direct-mail letters, the likelihood and size of donations increased by more than enough to cover the additional costs of the gifts. Bekkers and Crutzen (2007) found that envelopes with attractive color graphics were less successful than plainer envelopes, at least for continuing donors nurtured on the latter. Possibly graphics help when the campaign is designed to prospect for new donors, as Perrine and Heather

(2000) found that pictures on collection boxes placed in commercial establishments did increase donations.

One of the best studies, combining a laboratory and field experiment, is Vriens et al. (1998), in which they sent 16 different fundraising letters to 3000 different households per letter. Subjects were Dutch households that had contributed at least once in the past three years to a specific charity. They found that an optimal letter would contain no brochures or illustrations, not use bold print to amplify specific aspects, would contain a post-script summary, and would be signed by a Professor (this was a health research charity, and Professors are very highly respected in the Netherlands). This strategy would raise response rates by 9.8 percent and average donations by 5 percent. Response rates were extraordinarily high by American standards, and, considering the other cultural differences, it is unclear whether the specific conclusions would carry over to this side of the big pond.

Persuasive Strategies in Direct-mail Fundraising Letters

Because fundraising direct-mail letters are a distinct genre, with unique rhetorical moves and relational objectives, one would expect that persuasive messages might function differently in direct-mail fundraising letters than in other communication genres. Research has identified the functions of fundraising discourse (Bhatia, 1998; Connor, 1997), the rhetorical patterns (Abelen, Redeker & Thompson, 1993; Crismore, 1997; Lauer, 1997) that typify fundraising messages, as well as the social contexts (Bazerman, 1997; Myers, 1997) and cultural differences (Connor & Wagner, 1998; Graves 1997) that impact on fundraising communication.

The majority of this research is descriptive, identifying the rhetorical and linguistic features of the genre of direct-mail fundraising letters. Much of it is based on analysis of the fundraising texts housed in the Indiana Center for Intercultural Communication's (ICIC)

Fundraising Corpus. This Corpus codes nearly two million words of text from 900 fundraising documents collected from 236 nonprofit organizations, and is part of the American National Corpus. Within the Corpus are 245 direct-mail letters totalling 94,235 words.

Connor & Gladkov (2004) developed an operational system of persuasive appeals utilized in fundraising discourse, drawing from previous work by Connor and Lauer (1985) and rooted in the three appeal categories of Aristotelian persuasion: rational (*logos*), credibility (*ethos*) and affective/emotional (*pathos*). Rational arguments address the sensible, logical aspect of readers' minds. The most frequently employed tools to do so are "giving supporting evidence in the form of facts and statistics" and "showing how one event is the cause of another." Credibility appeals portray the writer and organization as trustworthy by "providing information to show first-hand experience or some authority on the subject" and "showing [the] writer's respect for [the] audience's interests and point of view." Affective appeals compel the audience to show empathy by "appealing to the Audience's views (emotional, attitudinal, and moral)." Rational appeals predominate in 48% of the Corpus letters, affective appeals in 28% and credibility appeals in 25%.

While all three persuasion categories are important, credibility appeals may be the most influential because of the importance of donor trust, which Handy (2000) identifies as the key challenge of nonprofits seeking to raise funds through direct-mail campaigns. Handy describes a variety of cues used by letter writers to build trust, including highlighting the charitable status of the organization, noting the longevity of the organization, using celebrity endorsement, or sharing with the reader the percentage of funds spent on administration. The Elaborated Likelihood Model (ELM) developed by Petty and Cacioppo (1981) provides further support for the efficacy of credibility appeals. ELM theory suggests that there are two routes to persuasion:

a central path, in which individuals are motivated to expend the energy required to cognitively process information and make an informed decision; and a peripheral path, in which individuals rely on short cuts. These researchers find that people are more likely to rely on credibility cues when they pursue the peripheral route, and more likely to rely on logical argument in the central route. To the degree that one would expect fundraising letters to be processed primarily through peripheral processing, one would expect credibility cues to be the persuasive appeals with the greatest impact on giving.

Visual Elements of Letter Layout

The relationship between persuasive strategy and giving is likely mediated by a wide range of factors, including elements of visual design. Conventional wisdom is full of suggestions about layout design, including the recommendation that direct-mail fundraising letters should use a variety of highlighting techniques to break the monotony of unbroken text, employ boldface or underlining to point out benefits, include a postscript, and make liberal use of numbered and bulleted lists (for summaries of conventional wisdom related to visual elements, see Smith, 1996; Vasilopolous, Goering & Nagelhout, 2004; Wheildon 2005). While some of these factors have been tested as individual strategies, no research has explored the interaction between physical layout and persuasive strategy.

We selected the use or non-use of bulleted lists for study. We chose this because of the disjuncture between conventional wisdom, which overwhelmingly endorses the use of lists, and practice. Only 19% of letters in the ICIC Fundraising Corpus contain bulleted lists. Consequently, we test whether bulleted lists result in higher donations, and whether there are interaction effects between the visual and persuasive treatments.

Linguistic Elements

A variety of linguistic elements related to direct mail fundraising are discussed in the conventional wisdom, including language choice and grammatical construction. In fact, the popular literature on how to write effective fundraising messages is full of advice about language use in fundraising letters (e.g., use clear, engaging, and personal language; "you" should be the most frequently used word in fundraising letters; use compact powerful words, colloquialisms, and familiar words—even clichés; avoid simile and metaphor, highly technical language, adjectives, first-person plural, big words, abbreviations, and foreign phrases) and about how to grammatically construct the fundraising letter (e.g., avoid semicolons, passive voice, and spelling errors; feel free to start sentences with a conjunctive; em dashes, ellipses, and contractions are all appropriate in fundraising letters).

This advice, when viewed in composite, endorses the notion that fundraising letters should be written at a relatively low level of readability. Readability scales, such as the Flesch Index, the Flesch-Kincaid Grade Level, the Fog Index, the Coleman-Liau Formula, or the Lix Formula, are indicators of the accessibility of a piece of writing. Generally, these formulas compute readability on the basis of sentence length, word length, and the general complexity of sentence structure. Although the conventional wisdom related to fundraising does not offer any specific advice related to the readability level that is best suited to this genre, the linguistic advice noted in the previous paragraph supports the claim that readability should be low. However, we found disjuncture between advice on the components of readability (endorsing the use of short sentences and paragraphs) and practice, with the majority of letters in the ICIC Fundraising Corpus targeting a 9th grade reading level. In addition, we observed inconsistency in reading level across the letters in the corpus (the average Flesch-Kincaid Grade Level of letters written by health and human services organizations was only 7.4; while the average grade level

of letters written by environmental organizations was 9.4). Thus, we compare the effects of letters with high vs. low readability on donations, alone and in interaction with the other treatments.

Methods

We employed the factorial survey method (also known as situation analysis, vignette analysis, or scenario analysis) to generate our data, an experimental design where respondents are asked to judge descriptions of varying situations (vignettes) presented to them.. We used a 3 (rhetoric) x 2 (linguistic) x 2 (visual) factorial design to generate 66 pairs of letters differing in at least one element of the design. Each letter was a fundraising message from a hypothetical university, named by drawing randomly from the eight compass directions (i.e., Southeastern University, Northern University, etc.). In each round of the study, subjects were given two letters to read and asked to imagine that they had decided to donate $100 divided across the two universities. As they read each letter, subjects were instructed to imagine that they were alumni from that university. Each subject followed this directive and allocated exactly $100. Subjects participated in 4 rounds, starting with a fresh $100 in each round. The letters were drawn from a rotation of the 66 pairs designed to assure that no subject received the same pair twice. For example, in round 1 the first subject might have been asked to look at a letter from "Southeastern University" (a mythical nonprofit educational institution) that includes rational appeals, high readability, and a bulleted list and another letter from "Northwestern University" (another mythical nonprofit educational institution) that utilizes credibility appeals, low readability, and a bulleted list. The process generated 164 observations of paired allocations. After four complete

rounds, subjects were asked a series of questions, inviting them to reflect on and explain their choices.

Although we initially employed more elaborate sampling techniques, low yields and budgetary constraints forced us to revert to a sample of convenience.[2] The sample consisted of 41 subjects who were alumni of IUPUI or current graduate students in Communication, Economics, or Philanthropic Studies.

To construct the 12 model letters presented to subjects, we first selected actual direct-mail appeals from the ICIC Fundraising Corpus that typified each of the three rhetorical treatments (rational, credibility, affective) and that contained lists in the body of the letter. We replaced the charity name with randomly-assigned university names, and produced two letters from each rhetorical treatment differing in their use of bullets. Finally, we produced two letters from each of the six rhetorical/visual model letters differing in their readability, as measured by the Flesch-Kincaid grade-level scale. We did so by modifying word length and sentence length, keeping the content and order of ideas as close as possible to the original corpus wording. It is of course possible that the particular way we reworded the letter had effects on giving for reasons unrelated to readability, but we are aware of no such differences. The high readability letters were written at a 12th grade level, with actual readability scores ranging from 11.7-12, and the low readability letters were written at a 6th grade level, with readability scores ranging from 5.9-6.2 on the Flesch-Kincaid Grade Level Index.

One limitation of this approach is that subjects were presented with a forced choice. They were asked to read a pair of letters, and then to definitely give a hypothetical $100 to one charity or the other or both. In the real world, the vast majority of direct-mail solicitation letters are thrown out unopened, and many of those that are opened do not convince the reader to make

a contribution. While opting for a forced choice design may limit the generalizability of our findings to real-world setting, it did facilitate testing the relative impact of persuasive strategies once a potential donor has made a decision to give, so perhaps our results are generalizable to the set of repeat donors. Finally, although our estimated rates of giving are higher than one would expect in the real-world, there is no obvious reason why forced choice should support response to one treatment over another, except, perhaps, with respect to the rhetorical dimension. It is possible that, say, an affective/emotional appeal is more successful in inducing subjects to consider donating whereas credibility is more important in forcing a split between recipients.

A second limitation inherent in factorial surveys is the hypothetical nature of choice. Although subject choices had no real consequences, there is no obvious reason why forced choices will be biased in any particular direction. Subjects desiring the approval of the experimenter may exaggerate their generosity, but the forced choice obviates the possibility of doing so. Another possibility is that because the choices are hypothetical, they reflect the subjects' first impulse rather than their carefully-considered opinions. However, the post-scenario explanations offered by subjects appeared generally thoughtful, suggesting that subjects (at least *ex poste*) exhibited realistic patterns of decisionmaking.

Finally, the factorial treatments were not completely crossed owing to early exhaustion of the project's budget. Appendix 1 shows the number of times each pair was actually tested. One pair was tested 7 times, as intended, but 3 pairs were never tested and the average pair (assuming order does not matter) was tested about 2.5 times. No clear pattern is evident in this table, so perhaps we are not far wrong in regarding the experimental treatments as random draws without replacement (for each subject) and with replacement (across subjects).

Results

Table 1 shows the distribution of amounts donated to the university represented by the first letter. This distribution was tri-modal, with 41% of observations going exclusively to one charity or the other, 22% splitting their donation evenly, and 37% splitting their donations unevenly. Those who gave all their money to the second charity were slightly more numerous than those who gave all their money to the first charity (23% vs. 18%). Possibly this represents a bias towards whatever is read last (which will be freshest in the subject=s mind when deciding how much to give to each charity). We report some evidence in favor of this interpretation below.

(Table 1 about here)

First, we present t-tests on those paired comparisons where there was a difference between characteristics (Table 2). There were 39 cases where a subject was presented with one affective and one rational letter in a round. Mean gifts to the charity represented by the affective letter were $47.31. Because gifts to one charity are identically $100 minus gifts to the other in any paired comparison, the proper statistical test is for a difference between mean gifts to one charity and a null hypothesis of $50 rather than for equality of the paired means. By this test, there was no statistically-significant difference in the impact of affective vs. rational letters. However, in the 51 cases where subjects chose between affective and credibility letters, mean gifts to affective were $40.39 and in the 35 cases of rational versus credibility, mean gifts to rational letters were $37.11. Both of these coefficients are significantly different from $50 at the 5% level. Thus, it appears that letters employing credibility strategies are most productive of donations, and there is little difference between the two runners up.

(Table 2 about here)

Readability matters. In the sample of high (H) vs. low (L) readability letters, mean gifts to H were \$67.83, allowing us to reject the null hypothesis at the .001 level. Visual style did not have a statistically-significant effect on giving. Mean gifts to schools using bulleted lists (B) were \$52.47 in the sample of all pairs that placed B against its alternative N (not bulleted).

In every case, subjects assigned greater donations to whichever letter was read second in each paired comparison. For example, in affective vs. rational comparisons, affective received \$46.55 when it was the first letter read and \$49.50 when it was the last. This is consistent with the expectation that whatever is freshest in the donor's mind is rewarded, but in no case was the difference in mean giving statistically significant.

In order to determine whether there are important interactions, round-specific effects, or subject-specific effects we turn next to a regression framework. The simplest approach would be to regress:

$$D1 = \alpha + \beta_1 A^1 + \beta_2 A^2 + \beta_3 R^1 + \beta_4 R^2 + \beta_5 H^1 + \beta_6 H^2 + \beta_7 N^1 + \beta_8 N^2 + \varepsilon$$

where:

D1 is donation to the first school in each paired comparison

A is a dummy variable equaling 1 if the letter has rhetorical characteristic A

R is a dummy variable equaling 1 if the letter has rhetorical characteristic R

H is a dummy variable equaling 1 if the letter has high readability

N is a dummy variable equaling 1 if the letter lacks bullet points

superscripts denote the first and second letter in a paired comparison.

But this approach seems unsatisfactory, as an example makes clear. Suppose that subject 1 always receives letters with rhetorical characteristic A (her letters differ in the visual or readability dimensions). Further suppose that subject 2 always receives letters with rhetorical

characteristic R. Then the regression would use the AR variation between subjects to estimate the effect of A even though we have no paired comparisons between A and R in the data. A=s effect would be hopelessly confounded with any individual-specific excluded variables.

Instead, we estimate models of the form:

(1) $D1_{it} = \alpha + \boldsymbol{\alpha_i} + \boldsymbol{\alpha_t} + \beta_1\, dA_{it} + \beta_2\, dR_{it} + \beta_3\, dH_{it} + \beta_4\, dN_{it} + \varepsilon_{it}$

where:

i indexes subjects and t indexes rounds.

D1 is donation to the school described in the first letter

dA is an indicator variable that equals 1 if the first member of the pair has characteristic A and the second does not, equals -1 if the second member has A and the first does not, and equals 0 otherwise. $dA = A^1 - A^2$

dR is an indicator variable equaling $R^1 - R^2$

dH is $H^1 - H^2$

dN is $N^1 - N^2$

$\boldsymbol{\alpha_i}$ is a vector of coefficients on 40 subject-specific dummy variables

$\boldsymbol{\alpha_t}$ is a vector of coefficients on 3 round-specific dummy variables

This approach exploits the panel nature of our data (subjects and rounds constitute the two dimensions of the panel). Because this specification will be unfamiliar to some readers, it is worth spelling out the underlying assumptions and the interpretation of coefficient estimates. The rhetorical dimension, with three possible characteristics (A, R, or C), is the most troublesome. If dA = 1, we know that the second letter possesses either characteristic R or C, so that either dR or dC (but not both) must equal -1; if dA = -1, either dR or dC must be +1. The coefficient on dA then indicates the effect on giving to the first school when R or C is substituted for A. If A has a positive effect on giving and dA = -1, the negative value of dA corresponds to

the fact that when A is in the second letter, giving to the first letter will be lower. One limitation of this approach is that it assumes that if both letters have the same rhetorical characteristic, it does not matter what value this characteristic takes.[3] Thus, an observation comparing, say, letters with characteristics ANH and ANL will have the same effect on regression estimates as an observation comparing RNH and RNL. This seems like a reasonable restriction, but it is a restriction.

Table 3 summarizes our results. Column 1 presents our preferred specification, in which individual-specific and round-specific effects were excluded but the error term was clustered by subject. This approach implicitly assumes that there are unobserved subject-specific random effects that are uncorrelated with the error term.[4] Rhetorical characteristics matter, with credibility producing higher levels of donations, but the statistical significance levels are lower than those in the simple t-tests reported above. Readability has a strong and highly significant effect on giving. Highly readable appeals, on average, result in donations that are $16.48 higher than less readable appeals. The point estimate of the effect of bulleted lists is small and statistically insignificant.

(Table 3 about here)

The remaining columns explore the robustness of these conclusions to alternative specifications. Column 2 includes interaction effects. Second-level interaction effects are the coefficients on variables constructed as the product of any pair of distinct characteristics (such as dH*dN).[5] Third-level interactions consist of the product of any three distinct characteristics, and under our design, there are two such nonredundant variables (dA*dH*dN and dR*dH*dN). The main effects are robust to inclusion of these addition variables, but it is difficult to interpret the full effects because none of the second-level interaction effects are statistically significant

whereas one of the two third-level interaction variables is significant. However, whether we look at the joint significance of all the second-level interaction variables, the third-level variables, or both the second- and third-level variables, we cannot reject the null hypothesis.

Column 3 replaces clustered standard errors with individual-specific fixed effects (the $\boldsymbol{\alpha}_i$ in (1). Results look just like those of column 1, except that the comparison favoring credibility over rational letters goes from marginally significant to marginally insignificant. Column 4 reports a median regression. Whereas OLS coefficients indicate the effect of an independent variable on the conditional mean of the dependent variable, median regression coefficients present the effect on the conditional median. This approach is useful when the error term is not classically distributed. Results are broadly similar to OLS, but all the independent variables except dN are statistically significant. This may be an artifact of our inability to correct for the clustering of errors without a larger data set (employing more rounds as well as more subjects). Column 5 reports another alternative that adjusts for the fact that the error term cannot be normal because gifts cannot be less than zero or more than $100. Double-tobit is commonly employed to deal with the resulting censorship bias (double because the dependent variable is constrained on both ends), and we are happy to employ it here. The estimates in column 5 are very close to those of column 1. We still prefer column 1 estimates because tobit is not robust to violations of the very strong assumptions the procedure makes, but it is comforting to know that accounting for censorship bias makes little difference in the present case.

Conclusions

Are there any universal rules governing the writing of direct-mail requests for donations, or is each campaign by each charity a special case? Fundraising texts and classical models of persuasion offer advice of both types, but the literature offers too few scientific tests of the

quality of that advice. We examine three dimensions of plausibly universal rules – rhetorical, linguistic, and visual, alone and in combination, and use factorial surveys to test the effect of these rules. However, we only begin the analysis here, because we look only at donations to institutions of higher education.

The first area explored was the rhetorical strategy used in constructing the persuasive argument. While previous research has tested the impact of particular persuasive tactics on giving (i.e., suggesting a particular donation amount, using the donor's name multiple times in the letter, framing outcomes positively versus negatively), no one has compared the impact of the three generic approaches to persuasion: logical, emotional, or credibility appeals. According to the results of this study, credibility appeals are the most productive of donations, with little difference observed between emotional and rational appeals. This finding is consistent with expectations from the Elaborated Likelihood Model (ELM). One would expect that the route to persuasion operative in fundraising is the "peripheral route," the path taken in situations where persuadees look for short cuts to reduce the cognitive energy required in processing persuasive appeals. ELM predicts that in situations such as this, credibility appeals are the most influential, and our results are in keeping with this expectation. In addition, our findings are consistent with the conceptualization of fundraising as relationship building, because credibility appeals are arguably the primary modes of persuasion in terms of establishing relationships.

The second area explored in this study was linguistic features of fundraising letters. Existing literature, mostly experience-based, provides few "meta-guidelines;" instead, the advice is typically offered as a seemingly random list of dos and don'ts. This study proffers "readability" as a useful meta-frame for examining language use in fundraising letters. The specific guidelines offered in popular sources related to fundraising (i.e., use short sentences and

short paragraphs, avoid big words) would construct a low-readability message, and this study tests whether that is indeed the most persuasive linguistic strategy in terms of soliciting donations. Our results indicate that in fact letters of higher readability are more successful in terms of donations. In part, this may be because subjects were told they were associated with the universities represented, and high readability is more important for unfamiliar charities. But the facts that the subjects were college-educated subjects and asked to donate to hypothetical universities suggest that this finding may not be generalizable. Regardless, our results imply that low readability is not a universally-valid criterion for direct-mail solicitation. Future research should examine other types of nonprofits that attract less-educated donors to see whether low readability is good advice anywhere.

The final element of written communication examined in this study was visual design. Use of bullet points is recommended in fundraising texts but rarely implemented in the representative letters contained in the fundraising corpus. This departure from recommended practice appears harmless, as we were unable to detect statistically significant effects of bullets on giving. Interestingly, in the follow-up interviews where subjects were asked to describe their decision-making process, the bullets were generally praised. Only two of the subjects mentioned that they did not like the bullets. Two did not mention the bullets at all, but all of the remaining subjects commented positively on the bullets, noting that the bullets made the letters "easier to read," "helped to draw attention to the kinds of things that support is needed for," and made the letter "friendlier."

While we were interested in the main effects in each of these three areas, we also were interested in the interactions among the three components. Regression analysis confirmed the

main effects for rhetorical strategy and readability but found no statistically-significant interaction effects.

These findings, limited as they are by the resource constraints described previously, do offer some indications about the communication strategies that may be most effective in direct-mail fundraising letters. Credibility appeals appear to be particularly effective, and communication theories suggest that this finding may be confirmed, by future researchers, to be a universal rule. In contrast, letters written at the 6th grade level are less effective than those written at the 12th grade level, a finding inconsistent with popular wisdom but likely particular to higher education appeals. Regardless of generalizability, we illustrate a methodology that is unfamiliar to many in our interdisciplinary field. The method has wide applicability for uncovering the determinants of giving and volunteering. Finally, factorial-survey experiments provide a relatively cheap and easy way for practicing fundraisers to refine their proprietary, campaign-specific testing.

Appendix 1 – Sample Size in Each Paired Treatment

	AHN	AHB	ALN	ALB	RHN	RHB	RLN	RLB	CHN	CHB	CLN	CLB
AHN	–	3	2	2	3	3	2	2	3	5	4	3
AHB		–	3	1	1	3	1	2	1	3	2	3
ALN			–	2	4	5	2	2	6	3	2	4
ALB				–	2	1	3	3	3	5	2	2
RHN					–	2	3	2	4	4	0	1
RHB						–	3	4	2	2	0	4
RLN							–	1	2	2	2	1
RLB								–	2	7	2	0
CHN									–	3	2	1
CHB										–	1	2
CLN											–	2
CLB												–

Notes

[1] These same disciplines study parallel questions regarding donations of blood, but the categories of persuasion are sufficiently different from those employed in the present study that we need not discuss blood donations studies further. For example, LaTour and Manrai (1989) combine in one variable (information provision) receipt of a brochure with vivid pictures, facts about the blood donation process and its consequences, and specific examples of need.

[2] In the initial protocol, we used two sampling frames provided by the Indiana University Foundation. The first frame consisted of frequent donors (who made donations of any size in 4 of the last 5 annual campaigns) who attended the Indianapolis campus (IUPUI) within the last 10 years and currently lived in the Indianapolis SMSA. The rare-givers frame was identical, except donations were made no more than once in the past 5 campaigns. Several waves of mailings produced only 22 subjects, so we approached on-site individuals to enlarge the sample.

[3]There are nine possible distinct pairs of rhetorical characteristics B AA, AR, AC, RA, RR, RC, CA, and CC. The two indicators for rhetorical characteristics can take only seven pairs of values, B (0, 0), (0, 1), (0, -1), (1, 0), (-1, 0), (1, -1), and (-1, 1). The indicator values map uniquely into rhetorical characteristics in most cases, but (0, 0) maps into AA, RR, or CC.

[4]Results clustering by both subject and round are nearly identical and not reported here.

[5]There is no need to construct an interaction variable equaling dH*dB because a list is either bulleted or not so the information in that variable is already contained in dH*dN. For the rhetorical dimension, we need interaction variables such as dA*dN and dR*dN; the third rhetorical interaction (dC*dN) is redundant. We also need an interaction variable obtained from multiplying dA by dR. The value of the coefficient on this variable allows us to test the assumption that the coefficient on dA is the same regardless of whether A changes because R changes or because C changes.

References

Abrahams, M.F. and Bell, R.A. (1994). Encouraging charitable contributions: An examination of three models of door-in-the-face compliance. *Communication Research*, 21, 131-153.

Abelen, E., G. Redeker, and S.A. Thompson. (1993). The rhetorical structure of U.S.-American and Dutch fund-raising letters. *Text,* 13, 323-350.

Andreoni, J. and Petrie, R. (2004). Public goods experiments without confidentiality: A glimpse into fundraising. *Journal of Public Economics,* 88, 16-5-23.

Basu, A.K., Basu, A. & Batra, R. (1995). Modeling the response pattern to direct marketing campaigns. *Journal of Marketing Research*, 32, 204-212.

Bazerman, C. (1997). Some informal comments on texts mediating fund-raising relationships: Cultural sites of affiliation. *Written discourse in philanthropic fund raising. Issues of language and rhetoric* (pp. 17-26). Indiana University Center on Philanthropy. Working Papers, 98-13. Indianapolis, IN.

Bekkers, R. and Crutzen, O. (2007). Just keep it simple: a field experiment on fundraising letters. *International Journal of Nonprofit and Voluntary Sector Marketing*, 12, 1-8.

Berger, P.D. & Smith, G.E. (1997). The effect of direct mail framing strategies and segmentation variables on university fundraising performance. *Journal of Direct Marketing,* 11, 30-43.

Bhatia, V. K. (1998). Generic patterns in fundraising discourse. *New Directions for Philanthropic Fundraising*, 22, 95-110.

Connor, U. (Ed.). (1997). *Written discourse in philanthropic fund raising. Issues of language and rhetoric.* IU Center on Philanthropy. Working Papers, 98-13, Indianapolis, IN.

Connor, U., & Gladkov, K. (2004). Rhetorical appeals in fundraising direct mail letters. In U. Connor & T. A. Upton (Eds.). *Discourse in the professions: Perspectives from corpus linguistics* (pp. 257-286). Philadelphia, PA: John Benjamins.

Connor, U., & Lauer, J. (1985). Understanding persuasive essay writing: Linguistic/rhetorical approach. *Text, 5*(4), 309-326.

Connor, U. & Wagner, L. (1998). Language use in grant proposals by nonprofits: Spanish and English. *New Directions for Philanthropic Fundraising: Understanding and Improving the Language of Fundraising*, 22, 59-73.

Crismore, A. (1997). Visual rhetoric in an Indiana University Foundation Annual Report. *Written discourse in philanthropic fund raising. Issues of language and rhetoric* (pp. 64-100). Indiana University Center on Philanthropy. Working Papers, 98-13. Indianapolis, IN.

Desmet, P. and Feinberg, F.M. (2003). Ask and ye shall receive: The effect of the appeals scale on consumers' donation behavior. *Journal of Economic Psychology*, 24, 349-376.

Diamond, W.D. & Gooding-Williams, S. (2002). Using advertising constructs and methods to understand direct mail fundraising appeals. *Nonprofit Management and Leadership, 12, 225-242.*

Falk, A. (2004). Charitable Giving as a Gift Exchange: Evidence from a Field Experiment. CESifo Working Paper No. 1218, Institute for the Study of Labor, Bonn, Germany.

Fraser, C., Hite, R.E., & Sauer, P.L. (1988). Increasing contributions in solicitation campaigns: The use of large and small anchor points. *Journal of Consumer Research*, 15, 284-287.

Graves, R. (1997). Dear friend (?): Culture and genre in American direct marketing letters. *Journal of Business Communications*, 34, 235-252.

Handy, F. (2000). How we beg: the analysis of direct mail appeals. *Nonprofit and Voluntary Sector Quarterly*. 29, 439-454.

Katzev, R.D. (1995). Applying *social psychology to charitable donations: Three experiments on nonprofit fundraising. Final report*. Washington, D.C.: The Aspen Institute Nonprofit Sector Research Fund.

LaTour, S.A. and Manrai, A.K. (1989). Interactive Impact of Informational and Normative Influence on Donations. *Journal of Marketing Research*, 26, 327-35.

Lauer, J. (1997). Fundraising letters. *Written discourse in philanthropic fund raising. Issues of language and rhetoric* (pp. 101-108). Indiana University Center on Philanthropy. Working Papers, 98-13. Indianapolis, IN.

Myers, G. (1997). Wednesday morning and the millennium: Notes on time in fund-raising texts. *Written discourse in philanthropic fund raising. Issues of language and rhetoric* (pp. 121-134). Indiana University Center on Philanthropy. Working Papers, 98-13. Indianapolis, IN.

Parsons, L. (2001). The Impact of Financial Information and Voluntary Disclosures on Contributions to Not-for-profit Organizations: A Field-based Experiment. Available at SSRN: http://ssrn.com/abstract=406760 or DOI: 10.2139/ssrn.406760.

Perrine, R.M. and Heather, S. (2000). Effects of picture and even-a-penny-will-help appeals on anonymous donations to charity. *Psychological Reports*, 86, 551-9.

Petty, R.E. and Cacioppo, J.T. (1981). *Attitude and persuasion: Classic and contemporary approaches.* Dubuque, IA: Wm. C. Brown.

Petty, R. E., Cacioppo, J. T., and Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement.” *Journal of Consumer Research, 10*, 135–146.

Reingen, P. (1982). Test of a list procedure for inducing compliance with a request to donate money. *Journal of Applied Psychology*. 67, 110-118.

Smith, G. (1996). Asking *Properly: The Art of Creative Fundraising*. London: The White Lion Press, Ltd.

Smith, G.E. and Berger, P.D. (1996). The Impact of Direct Marketing Appeals on Charitable Marketing Effectiveness. *Journal of the Academy of Marketing Science*, 24, 219-231.

Stone, R.A. (1992). The mass plaintiff: public interest law, direct mail fundraising and the. *Columbia Journal of Law and Social Problems*, 25, 197+.

Torre, L. & Bendixen, M. A. (1988). *Direct mail fund raising: Letters that work*. New York: Plenum Press.

Turner, T. & Yeakel, A.R. (1994). The University of Texas Law School Foundation: Small shop, big results. *Fund Raising Management*, 25, 24-28.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453-458.

Upton, T. A. (2002). Understanding direct mail letters as a genre. *International Journal of Corpus Linguistics, 7*(1), 65-85.

Vasilopolous, T., Goering, E.M. & Nagelhout, E. (2004, November). Development and implications of a taxonomy for fundraising letters. Paper presented at the Association for Research on nonprofit Organizations and Voluntary Action (ARNOVA), Los Angeles, CA.

Vriens, M., Scheer, H.R.V.D., Hoekstra, J.C., and Bult, J.R. (1998). Conjoint experiments for direct mail response optimization. *European Journal of Marketing* 32, 323-339.

Warren, P.E. and Walker, I. (1991). Empathy, effectiveness and donations to charity: Social psychology's contribution. *British Journal of Social Psychology*, 30, 325-337.

Weyant, J.M. & Smith, S.L. (1987). Getting more by asking for less: The effects of request size on donations of charity. *Journal of Applied Social Psychology*, 17, 392-400

Warwick, M. (2000). *The five strategies for fundraising success: A mission-based guide to achieving your goals*. San Francisco: Jossey-Bass.

Wheildon C. (2005). *Type and Layout*. Gainesville: The Worsley Press, FAP Books Inc.

Table 1: Amount Given to First Charity

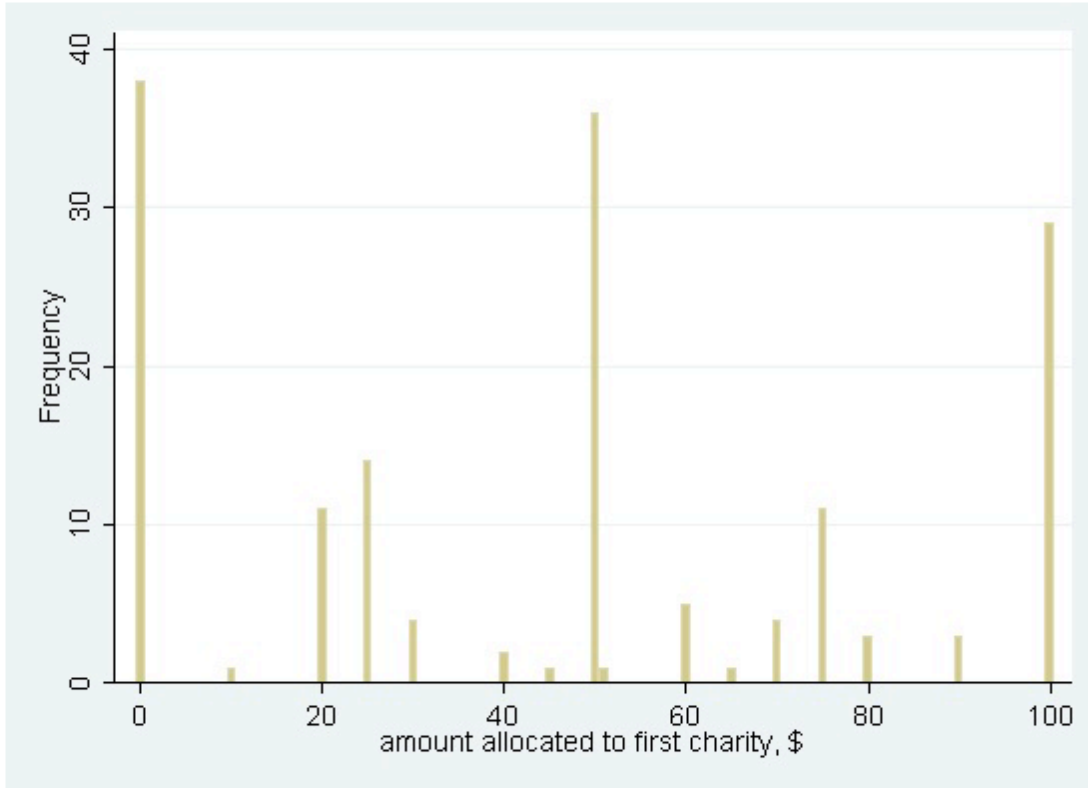

Table 2: Paired Comparisons

Comparison	**Mean (Standard Error)**	**N**
Affective vs. Rational (pooled)	A = $47.31 (5.83)	39
Affective first	A = $46.55 (6.67)	29
Rational first	A = $49.50 (12.53)	10
Affective vs. Credibility (pooled)	A = $40.39** (4.75)	51
Affective first	A = $36.43 (5.74)	35
Credibility first	A = $49.06 (8.29)	16
Rational vs. Credibility (pooled)	R = $37.11** (5.83)	35
Rational first	R = $32.11 (7.73)	19
Credibility first	R = $43.06 (8.89)	16
High vs. Low Readability (pooled)	H = $67.83*** (3.53)	92
High Readability first	H = $65.24 (5.63)	41
Low Readability first	H = $69.90 (4.51)	51
Bulleted vs. Not Bulleted (pooled)	B = $52.47 (3.81)	81
Bulleted first	B = $48.94 (4.80)	52
Not Bulleted first	B = $58.79 (6.21)	29

* denotes significance at the 10% level; ** at the 5% level; and *** at the 1% level.

Note: Significance for pooled estimates is for the difference between the mean and $50. * = 10% Significance for ordering estimates is for the difference between mean gifts to a characteristic when it is the first letter presented vs. the second, however none of these comparisons were significantly different from zero.

Table 3: Regression Results

	Preferred Specification[a]	Full Interactions[b]	Individual Effects[c]	Median Regression[d]	Double Tobit[e]
dA	-7.14 (5.10)	-5.07 (5.47)	-5.15 (5.00)	-20*** (5.03)	-7.27*
dR	-8.36* (4.67)	-7.12 (4.59)	-6.14 (4.80)	-20*** (5.16)	-8.47**
dH	16.48*** (3.98)	16.50*** (4.50)	14.79** (4.80)	25*** (4.11)	15.5***
dN	-1.98 (3.29)	-2.54 (3.17)	-3.42 (4.43)	0 (4.22)	-2.67
constant	47.64 (2.59)	53.46 (4.83)	34.68 (15.15)	50 (3.19)	
R^2	0.1647	0.2023	0.3527	0.1128	0.0236

Notes:

* denotes significant at the 10% level; ** at the 5% level, and *** at the 1% level.

dA indicates whether the two letters in a paired comparison differed in the Affective characteristic. Similarly for other variables, where R denotes rational, C denotes credibility, H denotes high readability, L denotes low readability, N denotes no bullets, and B denotes bulleted lists.

Standard errors in parentheses below parameter estimates. Significance levels are with respect to the excluded category. For the rhetorical dimension, there are three categories so we also tested for significant differences between the coefficients on dA and dR, which was never significant.

[a]Estimated by OLS with robust errors clustered on subjects.

[b]Interactions not displayed. None of the second-level interactions were individually or significant, and only the third-level interaction between R, H, and N was individually significant at the 10% level. The second-level, third-level, and combined joint significance levels did not reject the null hypothesis that interactions are unimportant. Estimated by OLS with clustered errors.

[c]Includes subject-specific individual effects. Estimated by OLS with robust errors. Individual effects were jointly significant at the 5% level. Results also including round-specific individual effects (not reported here) were similar but the round effects were insignificant.

[d]No individual effects or clustered errors. Pseudo-R2 reported.

[e]Dependent variable left-censored at $0 and right-censored at $100. This is the marginal unconditional effect on the dependent variable, with significance levels of the latent variable. Pseudo R^2 reported.