How to Ask for a Favor: A Case Study on the Success of Altruistic Requests

Tim Althoff*, **Cristian Danescu-Niculescu-Mizil**†, **Dan Jurafsky***
*Stanford University, †Max Planck Institute SWS
althoff|jurafsky@stanford.edu, cristian@mpi-sws.org

Abstract

Requests are at the core of many social media systems such as question & answer sites and online philanthropy communities. While the success of such requests is critical to the success of the community, the factors that lead community members to satisfy a request are largely unknown. Success of a request depends on factors like *who* is asking, *how* they are asking, *when* are they asking, and most critically *what* is being requested, ranging from small favors to substantial monetary donations. We present a case study of *altruistic* requests in an online community where all requests ask for the very same contribution and do not offer anything tangible in return, allowing us to disentangle what is requested from textual and social factors. Drawing from social psychology literature, we extract high-level social features from text that operationalize social relations between recipient and donor and demonstrate that these extracted relations are predictive of success. More specifically, we find that clearly communicating need through the narrative is essential and that linguistic indications of gratitude, evidentiality, and generalized reciprocity, as well as high status of the asker further increase the likelihood of success. Building on this understanding, we develop a model that can predict the success of unseen requests, significantly improving over several baselines. We link these findings to research in psychology on helping behavior, providing a basis for further analysis of success in social media systems.

1 Introduction

We live in a time where people increasingly turn to the web for help. Our needs, however, often go far beyond mere information from existing webpages and we need help from real people. For example, we ask for answers to specific questions on *StackOverflow.com*, for donations on *DonorsChoose.org*, or for help on online social communities such as *Reddit.com*. In each of these cases a user performs a request, which we define as an act of asking formally for something. All these communities rely heavily on their members to help satisfy the request. Yet, the factors that lead community members to satisfy a request remain largely unknown. Understanding the dynamics and factors of successful requests has the potential to substantially improve such communities by educating users about better formulating requests and promoting likely-to-succeed requests (Greenberg et al. 2013; Mitra and Gilbert 2014). In addition to these practical benefits, understanding the factors that make a request successful has implications for questions in social psychology and linguistic pragmatics.

Studies on the popular crowdfunding platform Kickstarter have shown that the success of a request depends most crucially on *what* is being requested, that is, whether it is a small favor like an answer to a simple question or a large financial contribution (Mitra and Gilbert 2014; Mollick 2014). Many other factors need to be controlled as well; what the giver receives *in return*, *when* they are asking, and even group dynamics, since people are more likely to give to projects that others are already giving to (Etter, Grossglauser, and Thiran 2013; Ceyhan, Shi, and Leskovec 2011; Mitra and Gilbert 2014). Satisfying a request on peer-to-peer lending or crowd-funding platforms can also bring a reward, and this also can drive the selection process. It is extremely difficult to disentangle the effects of all these factors in determining what makes people satisfy requests, and what makes them select some requests over others.

In this paper, we develop a framework for controlling for each of these potential confounds while studying the role of two aspects that characterize compelling requests: *social* factors (*who* is asking and how the recipient is related to the donor and community) and *linguistic* factors (*how* they are asking and what linguistic devices accompany successful requests). With the notable exception of Mitra and Gilbert (2014), the effect of language on the success of requests has largely been ignored thus far.[1]

Our goal is to understand what motivates people to give when they do not receive anything tangible in return. That is, we focus on the important special case of *altruistic* requests in which the giver receives no rewards. This controls for the incentive to obtain attractive rewards commonly offered on crowdfunding sites such as Kickstarter; the absence of external factors such as tangible rewards also makes the language itself all the more important in persuading others to help. In this domain we also do not need to consider crowdfunding-related marketing strategies such as emphasizing limited

[1] Linguistic factors have also been considered to influence the response quantity, quality, and speed to questions in online communities and social networks (Teevan, Morris, and Panovich 2011; Burke et al. 2007, inter alia).

time offers (scarcity) or showing that other people made the same decision already (social proof) (Cialdini 2001), which are known to manifest themselves in language (Mitra and Gilbert 2014). Second, we focus on requests that a single user can fulfill, thereby additionally eliminating group behavior effects such as herding (Ceyhan, Shi, and Leskovec 2011) or completing donation biases (Wash 2013). Finally, we focus on one community in which *what* is being asked for is held constant. This allows us to explore a large number of different requests of different individual users, at different times, that all have the same goal. Controlling for the request goal therefore allows us to study how to optimize a particular request solely by optimizing its presentation, and helps provide a direct practical benefit to the requester (by contrast, advising a requester who needs something to instead ask for something different may be advice of limited practical use).

We therefore chose to study donations in "Random Acts of Pizza", an online community devoted to giving away free pizza to strangers that ask for one. Random Acts of Pizza[2] (RAOP) is a community within the social news and entertainment website *Reddit.com*. Users can submit requests for free pizza and if their story is compelling enough a fellow user might decide to send them one, "*because... who doesn't like helping out a stranger? The purpose is to have fun, eat pizza and help each other out. Together, we aim to restore faith in humanity, one slice at a time.*[3]" A typical post might sound something like this: "*It's been a long time since my mother and I have had proper food. I've been struggling to find any kind of work so I can supplement my mom's social security... A real pizza would certainly lift our spirits* (Berman 2011)."

This platform addresses many of the potential confounds that complicate other platforms or studies: all requests ask for the same thing, a pizza, there are no additional incentives or rewards, each request is satisfied by a single user, users and requests are embedded in a social network within Reddit, and requests are largely textual. This dataset thus provides us with an unusually clear picture of the effect of language and social factors on success.

The remainder of this paper is organized as follows: inspired by studies in crowdfunding, user-to-user evaluations in social networks, and helping behavior in social psychology, we introduce a variety of textual and social factors that are potentially associated with successful requests. We use topic modeling and automatic detection to extract a particularly complex factor, the *narrative* structure of requests. We employ a logistic regression framework to test what factors matter in the community, showing that narratives are significantly correlated with success, and that signaling gratitude, the intention to reciprocate in the future, supporting the narrative with additional evidence, as well as a high status of the user within the community further increase the chance of success. We do not find any support for theories predicting that positive sentiment, politeness, and user similarity are associated with success. Thus, drawing from social psychology literature, our extracted high-level social features operationalize the relation between recipient and donor. We then demonstrate in a prediction task that the proposed model generalizes to unseen requests and significantly improves over several baselines.

[2] http://www.reddit.com/r/Random_Acts_Of_Pizza

[3] http://www.randomactsofpizza.com

2 The Dataset

Our dataset[4] contains the entire history of the Random Acts of Pizza Subreddit from December 8, 2010 to September 29, 2013 (21,577 posts total). To compute user features we further crawled the entire lifetime history of posts and comments across all Subreddits for all users involved in the RAOP Subreddit (1.87M submissions total). The community only publishes which *users* have given or received pizzas but not which *requests* were successful. In the case of successful users posting multiple times it is unclear which of the requests was actually successful. Therefore, we restrict our analysis to users with a single request for which we can be certain whether or not it was successful, leaving us with 5728 pizza requests. We split this dataset into development (70%) and test set (30%) such that both sets mirror the average success rate in our dataset of 24.6%. All features are developed on the development test only while the test set is used only once to evaluate the prediction accuracy of our proposed model on held-out data. For a small number of requests (379) we further observe the identity of the benefactor through a "thank you" post by the beneficiary after the successful request. This enables us to reason about the impact of user similarity on giving.[5]

3 Success Factors of Requests

Previous work on crowdfunding, helping behavior and user-to-user evaluations in social networks have pointed to a number of textual and social factors that could influence the success of a request.

3.1 Textual Factors of Success

Politeness A person experiencing gratitude is more likely to behave prosocially towards their benefactor and others (Tsang 2006; Bartlett and DeSteno 2006; McCullough et al. 2001). However, gratitude is only one component of politeness (Danescu-Niculescu-Mizil et al. 2013). Other indicators include deference, greetings, indirect language, apologizing and hedges. We ask a more general question: does a polite request make you more likely to be successful?

Evidentiality Some requests emphasize the *evidence* for the narrative or need. The literature on helping behavior literature suggests that urgent requests are met more frequently than non-urgent requests (Yinon and Dovrat 1987; Shotland and Stebbins 1983; Colaizzi, Williams, and Kayson 1984; Gore, Tobiasen, and Kayson 1997).

[4] Available at cs.stanford.edu/~althoff/raop-dataset/

[5] Reddit's front page showing the popular articles can skew exposure, but this will not effect RAOP as posts generally receive about two orders of magnitude less up-votes than would be necessary to appear on Reddit's front page.

Reciprocity In social psychology, *reciprocity* refers to responding to a positive action with another positive action. People are more likely to help if they have received help themselves (Wilke and Lanzetta 1970). Since in altruistic domains, there is no possibility of direct reciprocity, we hypothesize that recipients might pay the kindness *forward* to another community member, a concept known as "generalized reciprocity" (Willer et al. 2013; Gray, Ward, and Norton 2012; Plickert, Côté, and Wellman 2007). Feelings of gratitude can elicit this behavior (Gray, Ward, and Norton 2012). We hypothesize that the community would be more willing to fulfill the request of someone who is likely to contribute to the community later on.

Sentiment While many requests are fairly negative, talking about lost jobs, financial problems, or relationship breakups, some of them are positive, asking for pizza for birthday parties and other celebrations. Helping behavior literature predicts that positive mood is associated with a higher likelihood of giving (Forgas 1998; Milberg and Clark 1988). While these studies refer to the sentiment or emotional state of the benefactor, the most closely related linguistic feature that is available in this setting would be the sentiment of the text. Thus, the literature would predict that very positive requests are more likely to succeed. We additionally expect that very negative requests could be more successful, too, since they most likely describe very unfortunate situations of the requester.

Length Studies on the success of research grant proposals have shown that the simple factor of request length can be significantly related to funding success even when controlling for a variety of other factors (Lettice et al. 2012). We hypothesize that longer requests will be interpreted as showing more effort on the side of the requester and giving them the opportunity to provide more evidence for their situation.

3.2 Social Factors

Studies on crowdfunding have shown that the size of the social network of the project creator is associated with success (Mollick 2014; Mitra and Gilbert 2014). Work on user-to-user evaluations in online social networks suggests that the success of a request depends on who you are as a user and particularly that notions of user status and user similarity could be influential in the process (Anderson et al. 2012; Leskovec, Huttenlocher, and Kleinberg 2010; Guha et al. 2004). We study both status and similarity in this work.

Status Studies in social psychology have found that people of high status, e.g. defined by occupation or wealth, receive help more often (Solomon and Herman 1977; Goodman and Gareis 1993).

Similarity People are more likely to help those who resemble them (Colaizzi, Williams, and Kayson 1984; Chierco, Rosa, and Kayson 1982; Emswiller, Deaux, and Willits 1971). We predict that users will be more likely to give pizza to users who are like them in some way.

3.3 What Narratives Drive Success?

The textual part of a request, the narrative, has been shown to significantly influence the outcome in peer-to-peer lending platforms (Herzenstein, Sonenshein, and Dholakia 2011; Greenberg et al. 2013; Mitra and Gilbert 2014). In order to understand the nature and power of different narratives without coding them manually (Herzenstein, Sonenshein, and Dholakia 2011), we explore automatic methods of narrative extraction. Consider the following two pizza requests:

Example 1:

> "My gf and I have hit some hard times with her losing her job and then unemployment as well for being physically unable to perform her job due to various hand injuries as a server in a restuarant. She is currently petitioning to have unemployment reinstated due to medical reasons for being unable to perform her job, but until then things are really tight and ANYTHING would help us out right now.
>
> I've been both a giver and receiver in RAOP before and would certainly return the favor again when I am able to reciprocate. It took everything we have to pay rent today and some food would go a long ways towards making our next couple of days go by much better with some food."

Example 2:

> "My friend is coming in town for the weekend and my friends and i are so excited because we haven't seen him since junior high. we are going to a high school football game then to the dollar theater after and it would be so nice if someone fed us before we embarked :)"

While the first request (successful) goes into detail about hard times (and claims to reciprocate) the second one (unsuccessful) merely aims at "being fed".

To identify the different kinds of stories we draw on previous literature suggesting that narratives can be automatically extracted using topic modeling and related techniques (Chambers and Jurafsky 2009; Wallace 2012). We therefore perform topic modeling through non-negative matrix factorization (NMF) (Hoyer 2004) of a TF-IDF weighted bag-of-words representation (Salton and Buckley 1988) of the requests in our dataset. We additionally enforce sparsity on the topic distribution for each request to shape the topics in a way that captures most of a given request, and restrict ourselves to nouns (using the Stanford Part-Of-Speech Tagger[6]). We choose to use 10 topics and use a SVD-based initialization for NMF (Boutsidis and Gallopoulos 2008).

The resulting topics are shown in Table 1 along with descriptive names, the 15 highest-scoring terms and the success rate (fraction of requests that successfully obtained pizza). We observe that many topic clusters follow a specific theme and that their success rates vary dramatically (the average success rate is 24.6%). Topics MONEY1 and MONEY2 focus on money, and the high success rate of topic MONEY1 (32.3%) suggests that this is a particularly successful narrative. Topic JOB is similarly successful (31.9%) and features job related terms. A large number of requests further seem to come from college students talking about studying for classes and finals, their roommates, and the university (topic STUDENT). Another narrative in the data are requests

[6] http://nlp.stanford.edu/software/

Name	SR	Terms
MONEY1	32.3%	week ramen paycheck work couple rice check pizza grocery rent anyone favor someone bill money
MONEY2	23.6%	food money house bill rent stamp month today parent help pizza someone anything mom anyone
JOB	31.9%	job month rent year interview bill luck school pizza paycheck unemployment money ramen end check
FRIEND	17.0%	friend house night mine pizza birthday thing school site place family story way movie anything
STUDENT	23.2%	student college final pizza loan summer university money class meal year semester story kid school
TIME&FAMILY	23.5%	tonight night today tomorrow someone anyone friday dinner something account family bank anything home work
TIME	28.6%	day couple anything today work pizza help pay anyone home meal food ramen someone favor
GRATITUDE	27.0%	thanks advance guy reading anyone pizza anything story tonight help place everyone craving kind favor
STUDENT	23.2%	student college final pizza loan summer university money class meal year semester story kid school
PIZZA	20.0%	pizza craving hut story someone anyone domino money cheese thing request picture act title kind
GENERAL	24.1%	time pizza year people part work hour life thing lurker story anything someone month way

Table 1: Topics of requests identified by non-negative matrix factorization along with their success rate (SR). Note that the average success rate is 24.6%. Due to space limitations, we only display the 15 highest-scoring terms within each topic.

for and about family (topics TIME&FAMILY and MONEY2). This narrative can be identified by the usage of words indicating family relationships like kid, husband, family, mother, wife, and parents (not all of them are included in the top 15 terms). Topic FRIEND stands out since it noticeably worse than any other topic (17.0%). It captures requests asking for pizza for a friend that is in town, to cater for parties, or to provide culinary support for a night of movie watching with the girlfriend. We hypothesize that stories of this topic display little actual need for pizza (particularly compared to stories talking about money and job problems) and simply communicate a pizza craving by the requester. We further recognize that many requests employ previously defined factors such as gratitude ("thanks in advance" in topic GRATITUDE), providing pictures as additional evidence, and the intention to "pay it forward".

Automatic Narrative Detection This initial exploration suggests that narratives differ substantially in how successful they are. We define concise lexicons for each of the narratives to detect their presence or absence automatically. It would be possible to use a fully unsupervised approach using the lexicons generated through NMF but we find that the topic boundaries are often not very clear and would make it much harder to interpret the results. Instead, we use these topics as well as vocabulary from related LIWC categories (Linguistic Inquiry and Word Count; Tausczik and Pennebaker 2010) as inspiration to define concise lexicons for five different narratives identified through topic modeling (Money, Job, Student, Family, Craving). These lexicons along with example posts for each narrative are shown in Table 2. To measure the usage of these narratives in requests we define simple word count features that measure how often a given request mentions words from the narrative lexicons.

4 What Factors Are Predictive of Success?

In this section, we first introduce our methods for measuring each factor and then present results on which of them are predictive of success.

4.1 Measuring the Factors

Temporal Factors We control for temporal or seasonal effects in the data by measuring the specific months, weekdays, days of the month, or hour of the day as well as the the month of the request since the beginning of the community, i.e. the "community age".

Politeness We measure politeness by extracting all (19) individual features from the computational politeness model introduced by Danescu-Niculescu-Mizil et al. (2013).

Evidentiality We study a simple measure of evidentiality in RAOP posts: the presence of an image link within the request text detected by a regular expression. Many users provide evidence for their claims to be broke or injured by providing a picture, e.g., a screenshot of their empty bank account or a picture of their arm in a cast. Of the 84 images in a random subsample of about 2000 posts, 86% included some kind of evidence (an empty fridge, a job termination letter, the user themselves, etc.).

Reciprocity The concept of "paying kindness forward" to someone other than your benefactor after being the recipient of a kind action is referred to as *generalized reciprocity* in social psychology (Willer et al. 2013). We measure linguistic manifestations of generalized reciprocity by a simple binary feature based on regular expressions that indicates whether the text includes any phrases like "pay it forward," "pay it back" or "return the favor."

Sentiment We extract sentiment annotation for each sentence of the request using the Stanford CoreNLP Package[7] and encode whether a request employs an above average (median) fraction of positive sentiment sentences through a binary feature (same for negative sentiment). We further use count features based on lexicons of positive and negative words from LIWC (normalized by length) and a regular expression detecting emoticons to detect strong sentiment in text.

Length We use total number of words in the request as a measure of length and hypothesize that longer requests will be more successful on average.

[7] http://nlp.stanford.edu/software/

Narrative	Terms	Example Post
Money	money now broke week until time last day when today tonight paid next first night after tomorrow month while account before long Friday rent buy bank still bills bills ago cash due due soon past never paycheck check spent years poor till yesterday morning dollars financial hour bill evening credit budget loan bucks deposit dollar current payed	"*Broke until next paycheck*, Delaware. Really hungry and some pizza would be amazing right now. I had to pay to get my car repaired this *week*, leaving me with little *money until next Friday* when I get *paid* again. Some pizza would be really amazing. I would definitely pay it forward when I get *paid next week*."
Job	work job paycheck unemployment interview fired employment hired hire	"This is my first RAOP, low on money would really enjoy a pizza! Hey, my roommate and I are running low on cash. He lost his *job* last week and I had to pay his month's rent, and I'm going to have to until he finds another *job*. If someone could help us out with a pizza that would be great! Thanks!"
Student	college student school roommate studying university finals semester class study project dorm tuition	"*Studying* for *finals*, no time to go get food. Im *studying* for my last batch of *finals* before applying to *college* in the fall (transfer *student*, community *college* path). very hungry but being broke and having no calc textbook I'm really pressed for time :("
Family	family mom wife parents mother husband dad son daughter father parent mum	"Help out a *Dad* please? [...] I'm flat out broke until tomorrow with no food in the house for dinner tonight. My *daughter* is 2 and we usually do a pizza and movie night every once in a while, and she's been asking about it. I've got rent and car payment coming up, and bill collectors calling. I try to not let my *wife* know exactly how bad we are when it gets like this, but she mentioned we didn't have anything for dinner tonight, and I can't get groceries until tomorrow."
Craving	friend girlfriend craving birthday boyfriend celebrate party game games movie date drunk beer celebrating invited drinks crave wasted invite	"I went out with some *friends* earlier in the week and ended up lending my *friend* 20 bucks til he could get to an ATM. Long story short, we ended up pretty silly *drunk* and crashed at different houses so he never got a chance to pay me back. I get paid tomorrow and I could definitely tough it out, but I'd love to down a few slices before I spend the night cleaning up my apartment."

Table 2: The main narratives of textual requests discovered through topic modeling, the terms used to detect them automatically (sorted by frequency) and example posts illustrating each narrative.

Status We measure status in three ways. First, we use the *karma points* (up-votes minus down-votes) that Reddit counts on link submissions and comments, which define a notion of status in the Reddit community. Unfortunately, Reddit only publishes *current* karma scores for all users. Since we are only interested in features available at prediction time, we make sure to exclude all events that occurred after the request was first submitted and recompute karma scores from the user's full submission history up to that point in time. Second, we also measure whether or not a user has posted on RAOP before and thus could be considered a member of the sub-community. Third, we extract the user account age based on the hypothesis that "younger" accounts might be less trusted.

Narrative To measure the usage of all five narratives in requests we use word count features that measure how often a given requests mentions words from the previously defined narrative lexicons. We normalize these features by the total number of words in the request to remove length effects. Here, we use median-thresholded binary variables for easier interpretation but use decile-coded variants in the following prediction task.

4.2 Results

We model the success probability of a request in a logistic regression framework that allows us to reason about the significance of one factor given all the other factors using success as the dependent variable and the textual, social, and temporal features as independent variables.

The logistic regression results are summarized in Table 3 and are discussed next. We use a standard Likelihood Ratio test to compute significance scores.

Temporal Factors For temporal features we find that seasonalities within a day and a year did not differentiate significantly between successful and unsuccessful requests. However, the success rate decreases significantly with commu-

Coefficient	Estimate	SE
Intercept	-2.02^{***}	0.14
Community Age (Decile)	-0.13^{***}	0.01
First Half of Month (Binary)	0.22^{**}	0.08
Gratitude (Binary)	0.27^{**}	0.08
Including Image (Binary)	0.81^{***}	0.17
Reciprocity (Binary)	0.32^{**}	0.10
Strong Positive Sentiment (Binary)	0.14	0.08
Strong Negative Sentiment (Binary)	-0.07	0.08
Length (in 100 Words)	0.30^{***}	0.05
Karma (Decile)	0.13^{***}	0.02
Posted in RAOP before (Binary)	1.34^{***}	0.16
Narrative Craving (Binary)	-0.34^{***}	0.09
Narrative Family (Binary)	0.22^{*}	0.09
Narrative Job (Binary)	0.26^{**}	0.09
Narrative Money (Binary)	0.19^{**}	0.08
Narrative Student (Binary)	0.09	0.09

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3: Logistic regression results for textual, social, and temporal features displaying parameter estimates and standard errors (SE) for all features. Statistical significance is calculated in a Likelihood Ratio test. All features except the "student" narrative and sentiment significantly improve the model fit.

nity age. While the first 10% of requests had an average success rate of 36.6%, the last 10% were only successful in 16.9% of all cases (mostly, because the number of requests grew faster than the number of pizzas given out). Further, requests in the first half of the month tend to be more successful (26.4%) than in the second half (23.0%); we encode this as a binary feature.

Politeness Out of all 19 politeness features we find that only gratitude is significantly correlated with success when controlling for temporal, social, and other textual features. Thus, we only include gratitude in our final model.

Politeness is the expression of the speakers intention to mitigate face threats carried by certain face threatening acts toward another (Brown and Levinson 1987). We speculate that in this controlled environment there is very little room for face-threats as the roles are very well defined: requesters do not offend any potential giver by asking for a pizza but do exactly what is expected; potential givers can choose not to satisfy a particular request without face-threat (no direct interaction). Without face-threats there is no need for politeness to cover the acts of the requester or to provide the potential giver with a face-saving way to deny the request.

Evidentiality Including an image greatly increases the chance of success (second largest parameter estimate) providing strong support for the hypothesis that proving additional evidence makes you more likely to succeed. We attribute this to the fact that most pictures communicate need and urgency as well as establish an increased level of trust between requester and giver.

Reciprocity We find that the simple linguistic indication of willingness to give back to the community significantly increases the likelihood of success.

This finding raises the question whether those users who claim to give back to the community actually live up to this claim. To answer this question we restrict our data to only those requests that were actually successful. We find that on average 5.9% of successful users reciprocate (baseline rate). Out of those that claim to "pay it forward" after receiving a pizza 9.9% actually do. While this seems like a disappointingly small fraction it is conceivable that many users have not yet been able to help someone else out and might still do so in the future. And indeed, this fraction is significantly larger than the baseline rate according to a binomial test ($p < 0.01$).

Does gratitude predict generalized reciprocity as suggested by Gray, Ward, and Norton (2012)? We find that users that express gratitude in their request return the favor 7.2% of the time which exhibits only a slightly trend to be larger than the baseline (one-tailed binomial test, $p = 0.115$).

We also note that high status (karma in top 20%) is also positively correlated with reciprocity (one-tailed binomial test, $p < 0.05$).

Sentiment We find that sentiment stops being significantly correlated with success when controlling for the other variables. Similar results hold when using the fraction of sentences with positive/negative sentiment, thresholded versions of those features, other sentiment models and lexicons (LIWC) as well as emoticon detectors. This lack of relationship between sentiment and success may be a masking effect, due to the correlation between positive sentiment and other variables like reciprocity (Pearson correlation coefficient $r = .08$) and word length ($r = .10$).

Length Longer requests are significantly correlated with success. We attribute this to the fact that longer requests give the user the opportunity to provide more evidence for their situation. Length is arguably the most simple and accessible feature associated with success.

Status We find account age to be strongly correlated with karma ($r = .75$). This means that the "senior" users within the community tend to have high status (note that these are senior users that are still active as opposed to all senior user accounts). Therefore, we only include the karma score as a decile-coded variable (indicating the decile in the overall karma distribution) in the model. Status, in both in the Reddit community as well as the RAOP subcommunity, turns out to be strongly correlated with success. Having posted before in ROAP also has a particularly strong positive effect on success. People are more likely to help users that have contributed to the community in some form already (Willer 2009).

Narrative All narratives significantly improve the fit except for the "student" narrative. The "job", "money", and

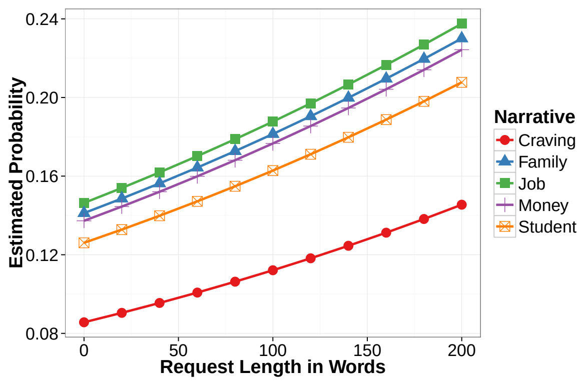

Figure 1: Estimated probability of success across request lengths for different narratives (top to bottom: Job, Family, Money, Student, Craving).

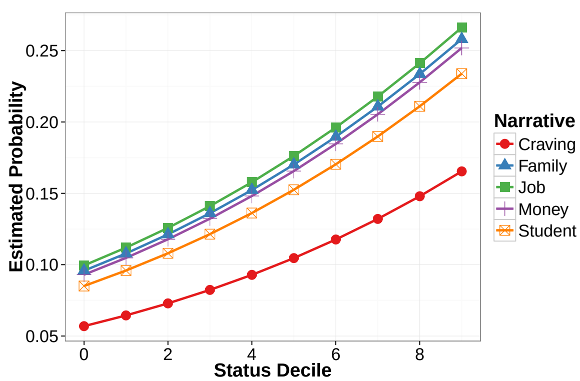

Figure 2: Estimated probability of success across status deciles for different narratives (top to bottom: Job, Family, Money, Student, Craving).

"family" narratives increase the predicted probability of success, while the "craving" narrative has strongly negative influence. This provides strong support for our hypothesis that narratives that clearly communicate need ("job", "money") are more successful than those that do not ("craving").

4.3 Interpretation

The logistic regression parameters correspond to changes in log odds space rather than probability space and are therefore more difficult to interpret. The change in probability space for different request lengths (median length is 74 words) is given in Figure 1 for the different narratives (assuming all other narratives are absent). Figure 2 depicts the estimated success probability for different values of status (karma). Both plots assume that the request does not include an image, gratitude, or reciprocity claim and assumes median values for karma (or length respectively) as well as community age (thus, the success probabilities are below average).

To understand the opportunity to optimize *how* one is asking for a favor and the importance to educate users about critical factors consider the following example: a short request (50 words) following the craving but no other narrative (assuming median karma and community age) has an estimated success probability of 9.8%. Using narratives that actually display more need, say the job and money narrative instead increases the chance to success to 19.4%, more than twice the previous probability. Now consider another user who is smarter about how she formulates her request. She puts in additional effort by writing more, say 150 words, and provides more evidence with a picture to support her narrative. She also makes sure to display gratitude to the community and offers to forward a pizza to someone else once she is in a better position. By tweaking her request in a simple way she increases her chances to 56.8%, a dramatic increase over the former request.

Features	ROC AUC
Random Baseline	0.500
Unigram Baseline	0.621^{***}
Bigram Baseline	0.618^{***}
Trigram Baseline	0.618^{***}
Text Features	0.625^{***}
Social Features	0.576^{***}
Temporal Features	0.579^{***}
Temporal + Social	0.638^{***}
Temporal + Social + Text	$\mathbf{0.669}^{***}$
Temporal + Social + Text + Unigram	$\mathbf{0.672}^{***}$

Table 4: Prediction results for logistic regression models using different sets of features. All models improve significantly upon the random baseline according to Mann–Whitney U tests ($p < 0.001$).

5 Is Success Predictable?

We demonstrated that textual, social and temporal factors all significantly improve the fit of a logistic regression model. Now, we study to what degree the model is able to generalize and predict the success of unseen requests from the held-out test set. Because of the unbalanced dataset and the trade-off between true and false positive rate associated with prediction we choose to evaluate using the area under the receiver operating characteristic (ROC) curve (AUC) which is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one.[8] Delong's test for two correlated ROC curves is used to test for statistical significant differences in the models (DeLong, DeLong, and Clarke-Pearson 1988).

Table 4 summarizes the performance of a L_1-penalized logistic regression model (Friedman, Hastie, and Tibshirani 2010) for different sets of features.[9] We include models using the standard uni-, bi- and trigram features (Mitra and Gilbert 2014) as baselines. We further include a random

[8]Note that this is closely related to the Mann–Whitney U statistic (Cortes and Mohri 2004).

[9]We also experimented with Support Vector Machines, with comparable performance.

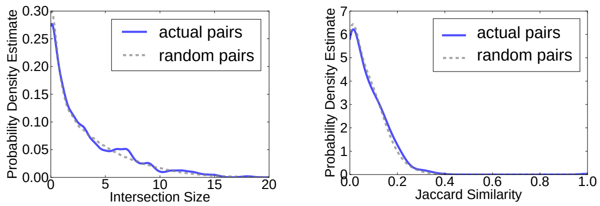

Figure 3: Kernel density estimate of the similarity distribution of actual and random pairs using Intersection Size (left) and Jaccard similarity (right) as similarity metric. We do not find any significant difference between the two distributions.

baseline for comparison of ROC AUC scores. It is important to note that there is no significant difference between our textual model with only 9 features and the uni-, bi- and trigram baselines which have orders of magnitude more features (Delong's test, $p \geq 0.612$). Both social and temporal features are predictive of success on held-out data as well. Combining the different feature sets yields significant performance improvements from temporal (0.579) over temporal + social (0.638; significant improvement over individual models at $p < 0.001$ in both cases), to all three sets of factors (0.669; significant improvement at $p < 0.001$ over textual model and at $p < 0.01$ over temporal + social model). Lastly, we demonstrate that a unigram model does not significantly improve predictive accuracy when combined with the proposed textual, social and temporal factors (0.672; Delong's test, $p = 0.348$). This shows that our very concise set of textual factors (narratives, evidentiality, gratitude, reciprocity, and length) accounts for almost all the variance that can be explained using simple textual models.

Although our best model (0.669) is far from perfect (1.000) all models significantly improve upon the random baseline (Mann–Whitney U test, $p < 0.001$). It is worth pointing out that we are purposely dealing with a very difficult setting — since the goal is to assist the user during request creation we do not use any factors that can only be observed later (e.g. responses, updates and comments), even though such factors have been shown to have strong predictive value (Etter, Grossglauser, and Thiran 2013; Mitra and Gilbert 2014).

6 Does User Similarity Increase Giving?

Social psychology literature suggests that individuals are more likely to help other individuals that are similar to themselves (Colaizzi, Williams, and Kayson 1984; Chierco, Rosa, and Kayson 1982; Emswiller, Deaux, and Willits 1971). To test this hypothesis we create a measure of user similarity by representing users by their interests in terms of the set of Subreddits in which they have posted at least once (prior to requesting pizza), and employing two different similarity metrics: intersection size between the set of the giver and receiver and the Jaccard similarity (intersection over union) of the two. Information about Subreddit overlap is easily accessible to users by a single click on a user's username. We compute the similarity of giver-receiver pairs and compare it with that of random pairs of other givers and receivers (pairs that did *not* occur in the data). The latter pairing is equivalent to a random rewiring of edges in the bipartite graph between givers and receivers which we use as a null model: if users are indeed more likely to give to other users that are similar to them we expect the similarity of actual giver-receiver pairs in our dataset to be significantly larger than the similarity of the randomly rewired pairs.

The resulting similarity distributions for both actual and random pairs for both metrics are shown in Figure 3 (kernel density estimate using a Gaussian kernel with bandwidth 0.5 for intersection size and 0.03 for Jaccard). The similarity distributions for actual and random pairs match very closely, a finding that is robust across both choices of similarity metrics. Thus, we conclude that we do not find any evidence that user similarity, at least in terms of their interest and activity as measured here, has a significant effect on giving. They may be other indicators of similarity including geography,[10] similar life situations or similar language use. In future work we plan to investigate other types of user similarity and their effect on helping behavior in online communities. Note that we do not include user similarity as a feature in the logistic regression model above since we only observe givers for a small subset of requests.

7 Conclusion

Online platforms have created a new mechanism for people to seek aid from other users. Many online communities such as question & answer sites and online philanthropy communities are created for the express purpose of facilitating this exchange of help. It is of critical importance to these communities that requests get addressed in an effective and efficient manner. This presents a clear opportunity to improve these online communities overall as well as improving the chance of success of individual requests. However, the factors that lead to requests being fulfilled are still largely unknown. We attribute this to the fact that the study of *how* one should ask for a favor is often complicated by large effects of *what* the requester is actually asking for. We have presented a case study of an online community where all requests ask for the very same contribution, a pizza, thereby naturally controlling for this effect and allowing us to disentangle what is requested from textual and social factors.

Drawing from social psychology literature we extract high-level social features from text that operationalize the relation between recipient and donor and demonstrate that these extracted relations are predictive of success. We show that we can detect key narratives automatically that have significant impact on the success of the request. We further demonstrate that linguistic indications of gratitude, evidentiality, and reciprocity, as well as the high status of the asker, all increase the likelihood of success, while neither politeness nor positive sentiment seem to be associated with success in our setting.

We link these findings to research in psychology on helping behavior (see Table 5). For example, our work extends

[10]We extracted location entities from the pizza requests but found them to be very sparse.

Category	Prediction by Literature	Finding of This Case Study
Gratitude	A person experiencing gratitude is more likely to behave prosocially towards their benefactor and others (Tsang 2006; Bartlett and DeSteno 2006; McCullough et al. 2001).	✔ Studies on gratitude typically focus on the gratitude experienced by the benefactor. We find that gratitude can be paid "forward" before the request becomes fulfilled and that expressions of gratitude by the requester significantly increase their chance of success. However, we find no evidence that politeness, more generally, has a statistically significant impact on success in our case study.
Reciprocity	People are more likely to help if they received help themselves (Wilke and Lanzetta 1970). The concept of paying kindness forward (rather than back) is known as "generalized reciprocity" in psychology (Willer et al. 2013; Gray, Ward, and Norton 2012; Plickert, Côté, and Wellman 2007).	✔ The language of reciprocity ("return the favor") is used in a variety of ways to signal the willingness to give back to the community by helping out another member in the future (generalized reciprocity). Such claims are significantly correlated with higher chances of success.
Urgency	Urgent requests are met more frequently than non-urgent requests (Yinon and Dovrat 1987; Shotland and Stebbins 1983; Colaizzi, Williams, and Kayson 1984; Gore, Tobiasen, and Kayson 1997).	✔ We find that narratives that clearly express need (job, money) are more likely to succeed that narratives that do not (craving). Additional support of such narratives through more evidence (images or text) further increased the chance of success.
Status	People of high status receive help more often (Solomon and Herman 1977; Goodman and Gareis 1993; Willer 2009).	✔ We find that Reddit users with higher status overall (higher karma) or higher status within the subcommunity (previous posts) are significantly more likely to receive help.
Mood/Sentiment	Positive mood improves the likelihood of helping and compliance (Forgas 1998; Milberg and Clark 1988).	✘ When controlling for other textual, social, and temporal factors we do not find the sentiment of the text (not the sentiment of the reader) to be significantly correlated with success.
Similarity	Persons are more likely to help other people when the similarity between them is high (Colaizzi, Williams, and Kayson 1984; Chierco, Rosa, and Kayson 1982; Emswiller, Deaux, and Willits 1971).	✘ Measuring similarity as the number of subcommunities that both receiver and giver are active in reveals no significant difference between actual pairs of giver and receiver and a null model of user similarity.

Table 5: A summary of predictions by literature on helping behavior in psychology compared to findings of this case study.

psychological results on offline communities to show that people behave pro-socially in online communities toward requestors who are of high status, display urgency, and who offer to pay it forward. Other novel contributions of our work include the finding that linguistic indications of gratitude lead to pro-social behavior, and the result that higher status users are more likely to demonstrate generalized reciprocity. Our results thus offer new directions for the understanding of pro-social behavior in communities in general, as well as providing a basis for further analysis of success in social media systems.

We must recognize a number of limitations: a shortcoming of any case study is that findings might be specific to the scenario at hand. While we have shown that particular linguistic and social factors differentiate between successful and unsuccessful requests we cannot claim a causal relationship between the proposed factors and success that would guarantee success. Furthermore, the set of success factors studied in this work is likely to be incomplete as well and excludes, for instance, group behavior dynamics. Despite these limitations, we hope that this work and the data we make available will provide a basis for further research on success factors and helping behavior in other online communities.

Acknowledgements We thank Christina Brandt and Jure Leskovec for many helpful discussions, Niloufar Salehi and Tuan Nguyen for helping with computing user similarity and statistical analysis, and Julian McAuley for the original idea to use the RAOP dataset. We also thank Reddit user "jimwll" who generously provided pizza when we most needed it and the anonymous reviewers for their altruistic comments. This work was supported in part by NSF IIS-1016909.

References

Anderson, A.; Huttenlocher, D.; Kleinberg, J.; and Leskovec, J. 2012. Effects of user similarity in social media. In *Proc. WSDM'12*.

Bartlett, M. Y., and DeSteno, D. 2006. Gratitude and prosocial behavior helping when it costs you. *Psychological science* 17(4):319–325.

Berman, J. 2011. Random Acts of Pizza: How to Donate. http://abcnews.go.com/US/random-acts-pizza-donate/story?id=13950694. Last accessed on 1/20/2014.

Boutsidis, C., and Gallopoulos, E. 2008. SVD based initialization: A head start for nonnegative matrix factorization. *Pattern Recognition* 41(4):1350–1362.

Brown, P., and Levinson, S. C. 1987. *Politeness: some universals in language usage*. Cambridge University Press.

Burke, M.; Joyce, E.; Kim, T.; Anand, V.; and Kraut, R. 2007. Introductions and Requests: Rhetorical Strategies that Elicit Response in Online Communities. In *Communities and Technologies 2007*. Springer. 21–39.

Ceyhan, S.; Shi, X.; and Leskovec, J. 2011. Dynamics of bidding in a P2P lending service: effects of herding and predicting loan success. In *Proc. WWW'11*.

Chambers, N., and Jurafsky, D. 2009. Unsupervised learning of narrative schemas and their participants. In *Proc. ACL'09*.

Chierco, S.; Rosa, C.; and Kayson, W. A. 1982. Effects of location, appearance, and monetary value on altruistic behavior. *Psychological Reports* 51(1):199–202.

Cialdini, R. B. 2001. Influence: Science and practice. *Boston: Allyn & Bacon*.

Colaizzi, A.; Williams, K. J.; and Kayson, W. A. 1984. When will people help? the effects of gender, urgency, and location on altruism. *Psychological reports* 55(1):139–142.

Cortes, C., and Mohri, M. 2004. Confidence Intervals for the Area Under the ROC Curve. In *Proc. NIPS'04*.

Danescu-Niculescu-Mizil, C.; Sudhof, M.; Jurafsky, D.; Leskovec, J.; and Potts, C. 2013. A computational approach to politeness with application to social factors. In *Proc. ACL'13*.

DeLong, E. R.; DeLong, D. M.; and Clarke-Pearson, D. L. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 837–845.

Emswiller, T.; Deaux, K.; and Willits, J. E. 1971. Similarity, sex, and requests for small favors. *Journal of Applied Social Psychology* 1(3):284–291.

Etter, V.; Grossglauser, M.; and Thiran, P. 2013. Launch Hard or Go Home! Predicting the Success of Kickstarter Campaigns. In *Proc. COSN'13*.

Forgas, J. P. 1998. Asking nicely? the effects of mood on responding to more or less polite requests. *Personality and Social Psychology Bulletin* 24(2):173–185.

Friedman, J.; Hastie, T.; and Tibshirani, R. 2010. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software* 33(1):1.

Goodman, M. D., and Gareis, K. C. 1993. The influence of status on decisions to help. *The Journal of Social Psychology* 133(1):23–31.

Gore, K. Y.; Tobiasen, M. A.; and Kayson, W. A. 1997. Effects of sex of caller, implied sexual orientation of caller, and urgency on altruistic response using the wrong number technique. *Psychological Reports* 80(3):927–930.

Gray, K.; Ward, A. F.; and Norton, M. I. 2012. Paying It Forward: Generalized Reciprocity and the Limits of Generosity. *Journal of Experimental Psychology*.

Greenberg, M. D.; Pardo, B.; Hariharan, K.; and Gerber, E. 2013. Crowdfunding support tools: predicting success & failure. In *Proc. CHI'13*.

Guha, R.; Kumar, R.; Raghavan, P.; and Tomkins, A. 2004. Propagation of trust and distrust. In *Proc. WWW'04*.

Herzenstein, M.; Sonenshein, S.; and Dholakia, U. M. 2011. Tell Me a Good Story and I May Lend You Money: The Role of Narratives in Peer-to-Peer Lending Decisions. *Journal of Marketing Research* 48(SPL):138–149.

Hoyer, P. O. 2004. Non-negative matrix factorization with sparseness constraints. *JMLR* 5:1457–1469.

Leskovec, J.; Huttenlocher, D.; and Kleinberg, J. 2010. Signed Networks in Social Media. In *Proc. CHI'10*.

Lettice, F.; Smart, P.; Baruch, Y.; and Johnson, M. 2012. Navigating the impact-innovation double hurdle: The case of a climate change research fund. *Research Policy* 41(6):1048–1057.

McCullough, M. E.; Kilpatrick, S. D.; Emmons, R. A.; and Larson, D. B. 2001. Is gratitude a moral affect? *Psychological bulletin* 127(2):249.

Milberg, S., and Clark, M. S. 1988. Moods and compliance. *British Journal of Social Psychology* 27(1):79–90.

Mitra, T., and Gilbert, E. 2014. The Language that Gets People to Give: Phrases that Predict Success on Kickstarter. In *Proc. CSCW'14*.

Mollick, E. 2014. The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing* 29(1):1–16.

Plickert, G.; Côté, R. R.; and Wellman, B. 2007. It's not who you know, it's how you know them: Who exchanges what with whom? *Social Networks* 29(3):405–429.

Salton, G., and Buckley, C. 1988. Term-weighting approaches in automatic text retrieval. *Information processing & management* 24(5):513–523.

Shotland, R. L., and Stebbins, C. A. 1983. Emergency and cost as determinants of helping behavior and the slow accumulation of social psychological knowledge. *Social psychology quarterly* 36–46.

Solomon, H., and Herman, L. 1977. Status symbols and prosocial behavior: The effect of the victim's car on helping. *The Journal of Psychology* 97(2):271–273.

Tausczik, Y. R., and Pennebaker, J. W. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 29(1):24–54.

Teevan, J.; Morris, M. R.; and Panovich, K. 2011. Factors Affecting Response Quantity, Quality, and Speed for Questions Asked Via Social Network Status Messages. In *Proc. ICWSM'11*.

Tsang, J.-A. 2006. Gratitude and prosocial behaviour: An experimental test of gratitude. *Cognition & Emotion* 20(1):138–148.

Wallace, B. 2012. Multiple narrative disentanglement: Unraveling infinite jest. In *Proc. NAACL HLT'12*.

Wash, R. 2013. The Value of Completing Crowdfunding Projects. In *Proc. ICWSM'13*.

Wilke, H., and Lanzetta, J. T. 1970. The obligation to help: The effects of amount of prior help on subsequent helping behavior. *Journal of Experimental Social Psychology* 6(4):488–493.

Willer, R.; Flynn, F. J.; Feinberg, M.; Mensching, O.; de Mello Ferreira, V. R.; Bianchi, A. M.; Choshen-Hillel, S.; Weisel, O.; Peng, K.; and Fetchenhauer, D. 2013. Do People Pay it Forward? Gratitude Fosters Generalized Reciprocity.

Willer, R. 2009. Groups reward individual sacrifice: The status solution to the collective action problem. *American Sociological Review* 74(1):23–43.

Yinon, Y., and Dovrat, M. 1987. The reciprocity-arousing potential of the requester's occupation, its status and the cost and urgency of the request as determinants of helping behavior. *Journal of Applied Social Psychology* 17(4):429–435.