

# Should We Use an Abstract Comic Form to Persuade? Experiments with Online Charitable Donation

ZIANG XIAO, University of Illinois at Urbana-Champaign, USA  
 PO-SHIUN HO, University of Illinois at Urbana-Champaign, USA  
 XINRAN WANG, University of Illinois at Urbana-Champaign, USA  
 KARRIE KARAHALIOS, University of Illinois at Urbana-Champaign, USA  
 HARI SUNDARAM, University of Illinois at Urbana-Champaign, USA

This paper examines the use of the abstract comic form for persuading online charitable donations. Persuading individuals to contribute to charitable causes online is hard and responses to the appeals are typically low; charitable donations share the structure of public goods dilemmas where the rewards are distant and non-exclusive. In this paper, we examine if comics in abstract form are more persuasive than in the plain text form. Drawing on a rich literature on comics, we synthesized a three-panel abstract comic to create our appeal. We conducted a between-subject study with 307 participants from Amazon Mechanical Turk on the use of abstract comic form to appeal for charitable donations. As part of our experimental procedure, we sought to persuade individuals to contribute to a real charity focused on Autism research with monetary costs. We compared the average amount of donation to the charity under three conditions: the plain text message, an abstract comic that includes the plain text, and an abstract comic that additionally includes the social proof. We use Bayesian modeling to analyze the results, motivated by model transparency and its use in small-sized studies. Our experiments reveal that the message in abstract comic form elicited significantly more donations than text form (medium to large effect size=0.59). Incorporating social proof in the abstract comic message did not show a significant effect. Our studies have design implications: non-profits and governmental agencies interested in alleviating public goods dilemmas that share a similar structure to our experiment (single-shot task, distant, non-exclusive reward) ought to consider including messages in the abstract comic form as part of their online fund-raising campaign.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; *Empirical studies in collaborative and social computing*.

Additional Key Words and Phrases: Abstract comics; Persuasion; Information framing; Social proof; Hierarchical Bayesian models

## ACM Reference Format:

Ziang Xiao, Po-Shiun Ho, Xinran Wang, Karrie Karahalios, and Hari Sundaram. 2019. Should We Use an Abstract Comic Form to Persuade? Experiments with Online Charitable Donation. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 75 (November 2019), 28 pages. <https://doi.org/10.1145/3359177>

Authors' addresses: Ziang Xiao, [zxiao5@illinois.edu](mailto:zxiao5@illinois.edu), University of Illinois at Urbana-Champaign, Computer Science, Urbana, IL, 61801, USA; Po-Shiun Ho, [pho11@illinois.edu](mailto:pho11@illinois.edu), University of Illinois at Urbana-Champaign, Computer Science, Urbana, IL, 61801, USA; Xinran Wang, [xinranw5@illinois.edu](mailto:xinranw5@illinois.edu), University of Illinois at Urbana-Champaign, Computer Science, Urbana, IL, 61801, USA; Karrie Karahalios, [kkarahal@illinois.edu](mailto:kkarahal@illinois.edu), University of Illinois at Urbana-Champaign, Computer Science, Urbana, IL, 61801, USA; Hari Sundaram, [hs1@illinois.edu](mailto:hs1@illinois.edu), University of Illinois at Urbana-Champaign, Computer Science, Urbana, IL, 61801, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2019 Association for Computing Machinery.

2573-0142/2019/11-ART75 \$15.00

<https://doi.org/10.1145/3359177>

## 1 INTRODUCTION

Are persuasive text messages more effective when expressed in comic form? Text messages abound, asking us to act, either towards personal wellness goals (e.g. in an exercise app: “you need 2,000 more steps to reach 10,000 steps for the day”), reminders (e.g. push notifications from a task tracking app: “pick up prescription from the pharmacy at 6PM”), or appeals from charities online (e.g. from a Wikipedia campaign on the Wikipedia homepage: “If everyone reading this donated \$5 our fundraiser would end today. Please donate to keep Wikipedia free.”). Making these short text messages more effective has important implications in stimulating behavior, especially for non-profits and governmental agencies interested in alleviating public goods dilemmas.

There is a rich history of prior work in psychology that has examined how variations in the construction of the text message alter decisions. In a seminal paper, Tversky and Kahneman [69] examined information framing: can we elicit different decisions, by describing the outcome either as a loss or a gain, even though the two messages are indistinguishable in terms of expected utility? They found that individuals are risk-seeking when confronted with certain loss and risk-averse in the presence of gains. Cialdini examined the introduction of social proofs in two different contexts: towel re-use in hotels [27], reduction of power consumption [64]. In the former study, replacing the standard text message in hotels asking the customer to re-use towels with the social proof (e.g. “75% of the people in this hotel have re-used their towels”) increased towel re-use. The latter study had an intriguing result—while introducing the social proof in the message *did not* reduce overall energy consumption, adding emoticons to the message with the social proof made a difference. Specifically, they added a smiley face ‘☺’ when the customer consumed less energy than average, and a frowny face ‘☹’ when their energy consumption was worse than average. By adding emoticons average energy consumption decreased. The Schultz et al. [64] study motivates us to examine the role of the comic form, a highly expressive and affective medium, in communicating messages intended to stimulate a specific behavior.

McCloud [45] defines comics as “juxtaposed pictorial and other images in deliberate sequence, intended to convey information and/or to produce an aesthetic response in the viewer.” The simple and humorous nature of a comic makes it a unique medium for delivering informative and memorable messages; the use of metaphor in comics can make the underlying meaning vivid and more memorable than using a straightforward description [46, 45]. There has been significant prior work in educational contexts: illustrating scientific facts [46], educating end-users about computer-security<sup>1</sup> [72], communicating with multi-lingual audiences [14, 51].

Despite the importance of the comic form in contemporary culture, and the use of the comic form in educational settings, there is surprisingly limited work examining the effects of the comic form to stimulate behavior. In particular, we focus on the use of an abstract comic form. By an abstract representation, (see Figure 1 for an example) we mean that the comic de-emphasizes character detail (face, eyes, etc.) or details about the locale. As McCloud [45] points out, using abstract representations for the comic allows the reader to project themselves onto the comic character. The abstract form has an additional benefit: since the form is visually spare, it allows for comic panel algorithmic synthesis and for the personalization of messages. In this paper, we used a three-panel comic in abstract form.

We use an online charitable donation task for our experiment: the task is a single-shot task, contributes to the public good with distant, non-exclusive rewards, occurs frequently, and easily tested for at scale. We chose a charity associated with Autism (Organization for Autism Research) for our experiment.

<sup>1</sup>Also used in introducing software: see [https://www.google.com/googlebooks/chrome/big\\_00.html](https://www.google.com/googlebooks/chrome/big_00.html) for an example.

Thus, in this work we answer two research questions in the context of the online charitable donation task:

**RQ1:** Does the use of the abstract comic form increase the level of donation over the plain text message?<sup>2</sup>

**RQ2:** What is the effect of introducing a social proof in the abstract comic, when compared to the comic without the social proof?

We would like to emphasize that it is unclear at the outset if a message in comic form ought to make any difference towards a charitable donation. An argument in favor is that comics communicate affect well and are an important element of our visual culture; an argument against is that from rational actor theory in classical Economics, since the comic includes the text message used in the text-only condition and thus cannot alter the expected utility, the comic ought to serve as an irrelevant factor. To the best of our knowledge, this is the first study to examine the persuasive power of an abstract comic in an online public good setting with distant and non-exclusive rewards.

We ran the experiment using Amazon Mechanical Turk ( $n = 307$ ; we paid each participant above minimum wage, at \$10.0/hr) where we randomly assigned each participant to one of the three conditions ('a plain text message', 'abstract comic', 'abstract comic with social proof'). The two comic conditions include the text used in the plain text condition. The social proof case adds one additional phrase to the comic that reveals the norm.

We performed a careful Bayesian analysis to analyze the data. We agree with the observations by Kay et al. [36], that beyond the impact on experimental replicability, shifting the question from "did it have an effect?" to "how strong was the effect?" is important to the HCI community. Two additional ideas—transparency, impact small- $n$  studies—motivate our use of Bayesian analysis. First, with a Bayesian model, the researcher foregrounds all the aspects of the model; there are no modeling assumptions that need checking, not already foregrounded in the model description. Thus the model and the results are amenable to further scrutiny. Second, a Bayesian model is valid at *every* value of  $n$  (the number of observations); we do not have to wait for  $n \geq 30$  to satisfy assumptions of Normality. For small  $n$  values, the result is of course affected by the choice of the prior; but by using weakly informative priors, we can ensure that the prior doesn't dominate inference. In a sense, as McElreath [47, Chapter 9] mentioned, Bayesian models make the *most conservative* inference given the data.

**Our findings:** We show that using the comic has a significant increase in donations over the plain text conditions, with a medium to large effect size of 0.59. Thus we can answer **RQ1** affirmatively. To answer **RQ2**, we show that while the comic with social proof increases the donation level over the comic without the norm, the effect size is very small (0.11) and the increase is not significant. To summarize, the comic form significantly increases donations over the plain text, but the presence of the social proof is not effective. We caution that the result holds for single-shot, public goods tasks. To fully understand the value of incorporating the social proof in the comic, for specific longitudinal tasks such as dieting, with distant, exclusive rewards and where habits can confound, we need future research.

**Design Implications:** The primary design implication of our findings: helping non-profits and governmental agencies with their online messaging strategies as they work to alleviate public goods dilemmas. In particular, public-goods dilemmas that require single-shot decisions (online charitable donations) are opportune candidates for intervention. We believe that these agencies

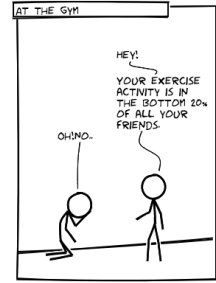


Fig. 1. An example single comic panel in abstract comic form.

<sup>2</sup>The abstract comic includes the same text from the plain text condition. Thus both conditions indicate the same expected utility.

can easily include the use of the comic form as part of their overall messaging strategy because the simplicity of the abstract comic form allows it to be easily synthesized (as we discuss in Section 6.2, towards the end of this paper) and to additionally incorporate social proofs.

The rest of this paper is organized as follows. In the next section, we discuss related work. Section 2 introduces two motivating ideas: the abstract comic form, and the idea of the social proof. Section 3 introduces our experimental method and in Section 4 we present our Bayesian analysis. In Section 5, we analyze the model for convergence and also discuss alternative models. Section 6 summarizes the main findings, discusses how we could algorithmically synthesize the comic panels and then identifies limitations. We present our conclusions in Section 7.

## 2 RELATED WORK

Now, we discuss related works in: 1) Using Comics to Communicate Ideas, 2) Persuasion through Visual Stimuli, 3) Persuasion for Public Goods, and 4) Social Proof in Persuasive Messages.

### 2.1 Using Comics to Communicate Ideas

McCloud [45] defines comics as “juxtaposed pictorial and other images in deliberate sequence, intended to convey information and/or to produce an aesthetic response in the viewer.” While reading comics book is commonly recognized as entertaining, comics have been examined as an effective way of communicating abstract and complex ideas to broad audiences [46, 14, 45, 72]. On one hand, the simple and humorous nature of comics makes it a unique media for delivering informative and memorable messages. By combining visual elements and text, comics makes the story more appealing. McDermott et al. [46] used comics to illustrate complex scientific facts. In education, comics has been used and examined as a useful media for reaching populations with various backgrounds [46, 14, 45]. Zhang-Kennedy et al. [72] created “Secure Comics” to educate end-users on computer security knowledge. On the other hand, the use of metaphor in comics can make the underlying meaning more vivid and memorable than using a straightforward description [46, 45]. Moreover, comics can contain a personal story which is incredibly powerful to create empathetic feelings, a key factor in persuasion, for readers [71]. Matsubara et al. [42] showed a link between comic’s contents and reader perceived emotions. Thus complex messages can be easily interpreted and memorized through the comic form. Although comics has shown its strength in communicating complex and memorable ideas, its utility in persuasion has not yet been fully explored. To the best of our knowledge, we are the first study looked at abstract comics in persuading people to cooperate in collective action dilemmas.

In our study, we chose abstract comics to persuade instead of other comic forms for the following reasons, 1) As the comic becomes more abstract, readers will be more likely to project themselves onto the character and internalizing the information that the character is trying to convey [45]. We believe such internalization gives an extra leg when the comic is designed to persuade. 2) The simplicity requires less cognitive load from the reader to consume the message, which is ideal for persuasion since persuadee’s attention is scarce [33]. 3) Comparing to other forms of the comic, the abstract comics contains least elements which lowers the cost to generate. The simplicity of the abstract comics allows us to explore the idea of algorithmically synthesizing persuasive messages into the comic form. Therefore, in this study, we examined the persuasive power of messages in abstract comic form.

### 2.2 Persuasion Through Visual Stimuli

Beyond the realm of textual forms, prior research showed that using visual representations, e.g., graphics, video, and comics, are attractive in persuasion. Selker et al. [65] used motivational graphics or memes from 9GAG to attract people’s attention and persuade people for energy conservation

behaviors. Consolvo et al. [18] visualized user's exercise data in the "Ubifit Garden" to persuade people to exercise. Sometimes, the use of visuals can evoke strong emotions that makes the message more persuasive. Iyer and Oldmeadow [32] found that people who saw the images of the Kenneth Bigley kidnapping were more engaged in the later civic campaign than those who read about the kidnapping via text. Visual stimuli were made for more memorable messages [53]. Visuals can influence decision making, especially when individuals make judgments using availability heuristics. Dey et al. [21] found that the video quality in the crowdfunding campaign plays a vital role in persuading people to support. Despite visual stimulus' benefits in persuasion, creating an effective visual persuasive message is often costly. The persuader needs to put time, effort and resource to create persuasive visual stimuli. In our study, we investigated the use of an abstract comic, a simple visual form that can be relatively easily created, for its persuasiveness in public goods dilemmas.

### 2.3 Persuasion for Public Goods

Public goods are non-exclusive and non-rivalrous, and thus encouraging individuals to participate in collective action dilemmas such as charitable donations is challenging [41, 31]. Therefore, external nudges play an important role in encouraging people to contribute. Researchers and policymakers have extensively studied public goods dilemmas to identify who contributes and how to persuade people to contribute [54, 5, 2, 51, 12, 57, 12]. Midden et al. [50] reported strong persuasive power for environmental sustainable behaviors when signaling personal goals in the persuasive application. Feiler et al. [24] found emphasizing altruistic reasons in a donation request can elicit more donations from people. Mankoff et al. [40] successfully used social technologies to leverage public commitment and competition in appealing to energy-saving behaviors.

A message composed with text is one of the most widely used methods to persuade individuals for behavior change or for charitable donations. Prior research indicates that textual messages are effective. Damgaard and Gravert [20] successfully used a simple email reminder with the decision deadline to elicit charitable donations. With a simple sign like "Turn off the tap when not in use," people reduced water consumption and engaged in energy conservation behaviors [48]. Vaish et al. [70] showed empathizing self-serving benefits can attract more people to sign up for a pro-social peer-to-peer service. Cotterill et al. [19] showed sending a pledge card with simple text "A list of everyone who donates a book will be displayed locally" encouraged 22 % more households to donate books for Children's Book Week. However, textual messages are less effective in catching persuadee's attention when compared to other visual forms, especially when persuadee's attention is limited. Moreover, when it comes to memorability, one of the key measures of persuasive effect, studies found that compared to graphics and video, text messages are the most difficult to recall [30, 63, 49].

### 2.4 Social Proof in Persuasive Messages

By "social proof", we refer to the idea that when individual's observation of either their friends or others they can relate adopted a behavior is persuasive for the individual to adopt the same behavior [16, 17]. "Social proof" is widely used in encouraging individuals to cooperate in collective action dilemmas [27, 64]. Goldstein et al. [27] conducted an experiment in a hotel on motivating environmental conservation. They found that descriptive norms in the persuasive message (e.g., "the majority of guests reuse their towels") were more persuasive than only mentioning environmental protection in the message. And the normative message got more effective when the statement was about a provincial norm (e.g., "the majority of guests in this room reuse their towels"). Amblee and Bui [1] studied "social proof" in online book reader communities and found that "electronic word

of mouth” affects book quality, author reputation, and book category popularity, which eventually influences people’s buying decision.

### 3 METHOD

In this study, we examined if the abstract comic can impact a decision with monetary consequence on Amazon Mechanical Turk. The persuasion task is to ask participants to make a charitable donation decision to the Organization for Autism Research with the real money.

In this section, we will first discuss the choice of task, then show how we construct persuasive messages (Section 3.3). Then, we introduce our experimental design (Section 3.2), discuss our study procedure (Section 3.4) and participant recruiting process (Section 3.5).

#### 3.1 Choice of Task: Charitable Donation

There are many different compelling behavioral contexts on which to test the role of the comic form: personal wellness goals (e.g., diet, exercise), mundane tasks (e.g. “pick up dry cleaning”), as well as broader public-goods issues (e.g. “take the flu shot;” “donate to cure cancer”).

We identify four criteria that consistent with prior literature [39, 68, 62, 60] to guide us, in our choice of the experimental behavioral context: nature of the reward; single shot tasks; an ecologically valid task; absence of specialized knowledge to perform the task. First, we would like the rewards to be distant, and non-exclusive, rather than proximal and exclusive so that individuals won’t perform the task in anticipation of the immediate reward. Thus public goods dilemmas (e.g. “reducing carbon footprint;” “taking the flu shot;” “contributing to public knowledge”) are all candidates. Second, while some longitudinal tasks (e.g., losing weight; eating healthy) have distant rewards (losing weight, or maintaining a diet takes time), and can positively affect the public good (with more healthy people, in the long-run, insurance rates will fall), these tasks are prone to habit formation, a potential confound. Furthermore, single-shot tasks such as “pick up yogurt at the grocery store today”, often prompted by text reminders from our calendars or task-tracking apps, have an immediate, exclusive reward. Third, we would like to ensure that the experimental task is ecologically valid—a task that these individuals would be actually asked to perform in the real world, outside of the experimental context. Fourth, we would like the task not to require specialized knowledge (e.g. “asking doctors to make a decision”), so that other researchers could easily replicate and scale.

Online charitable donation tasks satisfy those criteria as 1) they are single-shot tasks, 2) they contribute to the public good with distant, non-exclusive rewards, 3) requests for charitable donations frequently occur online, and 4) the task allows for experimental replication.

We also considered two other in-the-wild experimental scenarios: buying healthy food and exercise. We considered the task of persuading individuals to buy healthy food by sending them push notifications on the smartphone when they were shopping for groceries. We encountered two challenges. First, we could not control for confounds in the decision making (if there are discounts on salad / healthy items; if the shopper was in a rush; if the shopper was shopping with friends who were pursuing a healthy lifestyle). Second, the reward had a possible confound for some participants, healthier food may make them feel better, and thus they may make the choice to buy healthier food not because of the message, but because of the anticipated reward. We also considered persuading individuals to adopt healthy behaviors, such as taking a walk or going to the gym. Then, we can measure gym visits or exercise outcomes. However, the central challenge here is that those activities are prone to habit formation, a confound. If habit forms, it may be more important in causing repeated behavior than the message received.



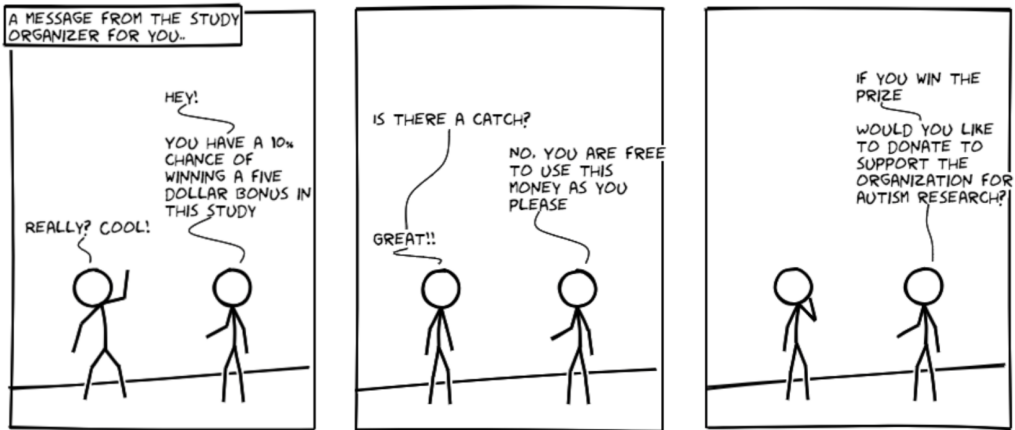


Fig. 2. Messages in the abstract comic form. Same as the text messages, the three-panel comic strip communicates three points: 1) Participants will have 10% of a chance winning \$5 bonus upon the completion of the study (see the first panel); 2) Participants are free to use the money as they please (see the second panel); 3) Participants can donate this bonus to the Organization for Autism Research (OAR) (see the third panel).

In our study, we choose the Organization for Autism Research (OAR) for three reasons. First, Autism Spectrum Disorder (ASD), is a well-recognized developmental disorder that impairs communication and behavior [4]; ASD provides basic interest for the participants to support the related charitable organization. Second, the Organization for Autism Research is one of the most visible ASD related organizations that helps individuals with autism and provides assistance to their parents, families, teachers, and caregivers. The goal of OAR is clear and reputable so participants won't question the authenticity of our message's motive. Finally, we wished to avoid a charity associated with a life-threatening condition such as cancer as it may create an experimental confound: we don't know if someone donates because their intrinsic desire to help with a life-threatening situation. While ASD can have serious consequences on the well being of those who have it, the public perception is that ASD is not-life threatening [4].

### 3.2 Experiment Design

Since the main goal for this study is to compare the power of a persuasive message in two forms, the abstract comics and the text, we first constructed two experimental conditions, the comic condition (see Figure 2) and the text condition. In the comic condition, participants will read a message asking if they are willing to support a charity in a three-panel abstract comic strip, whereas in the text condition, participants will receive the same message in plain text form.

To test the idea of social proof, we then added a third condition, the comic with the social proof condition. Participants will read a three-panel comic strip that has the same content as the comic condition but added one line text indicating the normative behavior (see Figure 3). To gather the basic statistics to create social proof, we first ran a pilot study with the first two conditions ( $n = 60$ ) on Amazon Mechanical Turk and used the donation statistics as part of the social proof message. In the pilot study, 87% of the participants donated a non-zero amount.

The objectives of the persuasive message in all three conditions are the same: persuading participants to donate to a charity from his/her own pocket. Similar to Lee and Hsieh [39] and Saunders et al. [62], the money participants will use is part of their study reward, a prospective

bonus reward (10% chance of winning \$5 bonus). We randomly assigned study participants to one of the three conditions; the participants are free to make a decision on the amount of donation, including not donating at all.

### 3.3 Constructing Persuasive Messages

Now we discuss construction of the three different messages: a plain text message, a comic, and a comic with social proof. Each message communicates three points: 1) Participants will have 10 % of a chance winning \$ 5 bonus upon the completion of the study. 2) Participants are free to use the money as they please. and 3) Participants can donate this bonus to the Organization for Autism Research (OAR). Therefore, in the text condition, study participants will read the following message,

*You have a 10% chance of winning a five dollar bonus in this study. You are free to use this money as you please. If you win the prize, would you like to donate to support the Organization for Autism Research?*

In the two comic conditions, we created three-panel comic strips to communicate the *same* text message. We used three panels to communicate each of the three points in the message.

While aesthetic considerations govern the choice of the number of panels in the comic form, two factors influenced our choice of the number of panels: the number of points in the message, and communicating the idea of a conversation, over time, between two individuals. Since we wish to communicate three points (chance of winning bonus; free to use the money as they please; voluntary donation) in the short comic, to ensure that each point was salient, we chose to communicate each point in a different panel. Comic panels along with the text bubbles within a panel fragment time and space. Time flows in two ways: vertically in each panel through the dialogue between the characters and across panels. Since this is a short conversation, three panels seemed appropriate. While adding more panels would elongate the perceived sense of time, we wished to be economical in our communication of the interaction between two characters. We can control of the number of panels in our comic generator (see Section 6), and we leave the question of how the number of panels (and thus the perceived sense of time) affects message comprehension for future work.

To help the reader connect across panels to achieve closure McCloud [45, Chapter 3] suggests that it is essential to match the panels. Different techniques to match exist, including “moment-to-moment”, “action-to-action”, “scene-to-scene” etc. help the reader bridge different time scales (see [45, p. 71] for a summary). We used the “action-to-action” matching technique [45] since all the activity occurs in one scene, and where the activity depicts a brief dialogue between two characters. In particular, we matched the panels on the gesture of the first character (the message recipient), while retaining a neutral gesture for the second character (who delivers the message).

In the comic with social proof condition, we created social-proof by adding one sentence on the last comic panel indicates the percentage of people in our study donated (Figure 3). The percentage (87%) corresponds to the number of people who donated a non-zero amount in the pilot study.

For the purposes of this study, we created a comic generator to generate comic strips used in the study. The generator leveraged an open source comic library, “cmx.io” [28], to generate comic figures and used “rough.js” [66] to generate other elements such as text bubbles and outline frames. The generator impersonates the style found on “XKCD” [52] comic. Our focus is not the XKCD style, but the fact that the generated comic is abstract. We believe the generator has the potential to be further developed as a general framework to automatically synthesize pure-textual persuasive messages into abstract comic forms (we discuss this aspect further in section Section 6).



### 3.4 Study Procedure

Once participants consented to join the study, we ask them if they are familiar with the Autism Spectrum Disorder (ASD). Then, each participant watched a short video produced by the Organization for Autism Research that promotes its fundraising activity "RUN FOR AUTISM" [59]. After watching the video, we asked participants to summarize the video using free text and ask them to provide their opinion about the effectiveness of the video. The recruiting message specifically mentioned the task is soliciting their opinion on video message's effectiveness. There are two main goals for this part of the study: first, similar to the informational materials (e.g. text) used in other donation studies [39, 43, 24], the video provides context. We want to make sure prior familiarity with autism do not confound our study; second, we want our main task less intrusive as soliciting charitable donation is not a common task on Amazon Mechanical Turk. We did consider establishing context by other means, such as an informative page with text and images on the organization of autism research. We decided against such an approach, because we could not guarantee that all subjects would have read the page. Furthermore, making subjects answer questions about the web-page to check if they had read the page carefully, seemed artificial, impinging on ecological validity. Finally, the linearity of the video is an advantage: subjects could advance to the next page only after watching the video, ensuring consistent knowledge to the best of our ability.

We then randomly assigned participants to one of the three conditions and asked them to read the corresponding persuasive message (text, three-panel comic, three-panel comic with social proof). In the message, we provided the participant with a 10% chance of winning \$5 additional bonus reward. We also provided them with the opportunity to donate to the Organization for Autism Research (OAR) which is the charity mentioned in the video they previously watched.

To best demonstrate the persuasiveness of the message, in our messages, we made it clear to participants that they can use the bonus freely as they please. The donation is completely voluntary and irrelevant to the compensation. The default donation amount for everyone is set to \$0 to control for the default effect. Similar to Lee and Hsieh [39], we also diffused the responsibility of donation amount among all participants. Before the participants make their decision, they read "The total amount of money allocated to [the charity] by all the winning participants will be aggregated and donated at the end of the study." Then, we asked the participants to decide the amount of money they are willing to donate on a slider bar with \$0 and \$5 as two extreme ends. To increase the realism of our study, the donation decision is not hypothetical: the winning participants receive the amounts that they choose keep; as study organizers we donate to the Organization for Autism Research (OAR) the cumulative sum of the participant donations.

Before leaving the study, participants need to filled a demographic questionnaire about their gender, income and education; participants had the option of declining to state an answer for each question.

At the end of the study, we randomly chose 10% of the participants, donated to OAR based on their decision, and rewarded them with part of the bonus that they wished to keep.

### 3.5 Participant Recruitment

In this study, we recruited our study participants from Amazon Mechanical Turk. Although the Amazon Mechanical Turk, a crowdsourcing platform, has been widely used to gather human intelligence in AI research and social science experiments [55, 7, 11, 10, 39, 62, 3, 68], we should be cautious when using such platform as the participant selection criteria is not transparent [38, 56]. In our study, we chose the Amazon Mechanical Turk for the following reasons. First, our persuasion task is about online charitable donation which targets internet users. Second, crowdsourcing platforms will help us reach a more diverse sample than using the researchers' own

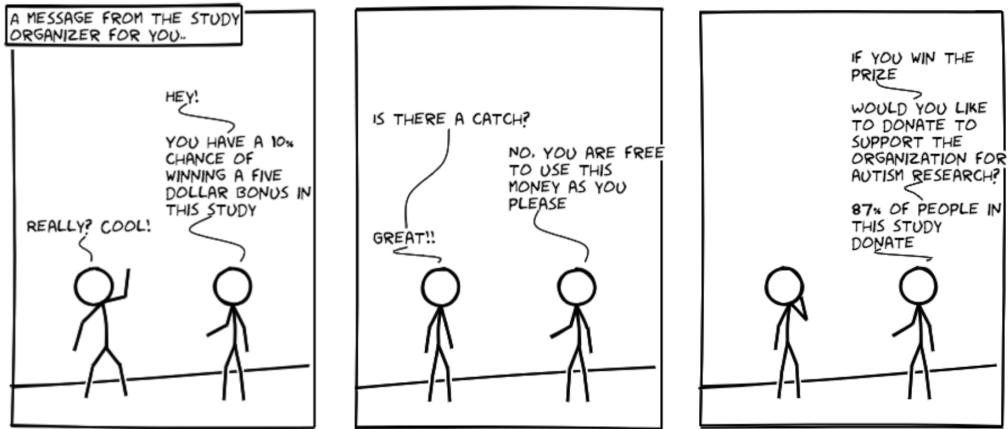


Fig. 3. Messages with social proof. In addition to the three points that we wish to communicate (chance of winning bonus; free to use the money as they please; voluntary donation), the comic with social proof communicates the idea of the social proof at the third comic panel. The figure “87%” comes from our pilot study

social network to attract participants [11]. Third, the main motive for Amazon Mechanical Turk workers is monetary rewards [7]. In our study design, Amazon Mechanical Turk subjects may be more sensitive to the monetary reward they receive, making them less amenable to persuasion. Finally, the use of Amazon Mechanical Turk subject pool by multiple studies [10, 39, 62, 3, 68] on charitable donation decisions informed our decision to use the Amazon Mechanical Turk subject pool.

Motivated by studies that show that populations on Amazon Mechanical Turk are diverse and mirror the US population [11, 6, 7], we recruited our participants from Amazon Mechanical Turk. However, we are aware of research that raises concerns in Amazon Mechanical Turk’s sample representativeness [38, 56]. One potential solution is to use a panel company’s population (e.g., Qualtrics). However, this method also has concerns in that the researcher can not directly cross-validate the sample’s representativeness—we have to trust the company’s assertion.

We published our HITs on Amazon Mechanical Turk titled “A short survey about communicating autism campaign ads”. The compensation was \$10/hr, and the workers would get the compensation regardless of their performance. To ensure quality, the HIT is limited to English Speakers in US and people who have a 95% Approval Rate. On the HIT page, we instructed that repeated responses would be rejected. We told the participants that they would see a link to our experiment site (Qualtrics).

## 4 RESULTS

In this section, we will first present the raw data used in the analysis, then introduce the Bayesian Model we used for data analysis and our analysis result.

### 4.1 Raw Data

In total, we have 307 participants joined our study, 101 participants received the message in the text form, 102 participants received the same message in the abstract comic form, and 104 participants received the abstract comic message with social-proof. We ran an outlier analysis using Tukey

Fence on participants' completion time and removed 30 participants from our dataset. The following analysis is based on a dataset with 277 participants, 94 of them is in pure-text condition, 91 of them is in comic condition, and 92 of them is in comic-social-proof condition. Of those 277 participants, 150 were self-identified as male, and 126 were self-identified as female, 1 participant chose not to disclose. The median annual household income of our study participants is \$50,000 - \$59,999, which is close to the median annual household income in US (\$ 57,652).<sup>3</sup> 71 % of study participants earned at least college degree. All participants were familiar with the Autism Spectrum Disorder and acknowledged the importance of autism research.

Figure 4 compares the distributions of the charitable donations across the three conditions. Among all 277 participants, 223 (80.5%) participants donated non-zero amount to support the autism research; 67 (71.3%) participants from the text condition, 75 (82.4%) participants from the comic condition, and 81 (88.0%) participants from the comic with social proof condition.

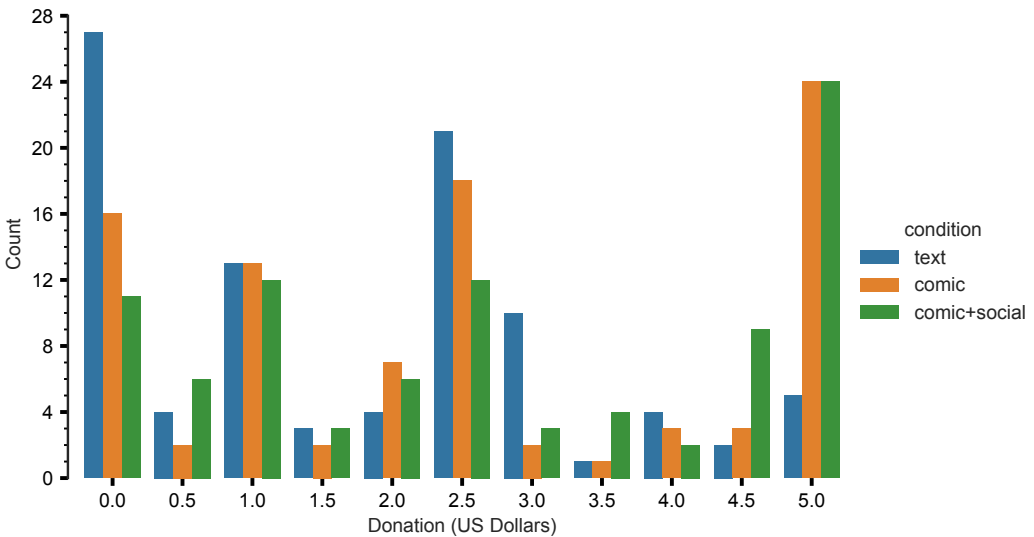


Fig. 4. The distribution of the amount of money participants decide to donate to the charity in each of the three conditions: text only; comic; comic with social proof. Among all 307 participants, 223 (80.5%) participants donated non-zero amount to support the autism research; 67 (71.3%) participants from the text condition, 75 (82.4%) participants from the comic condition, and 81 (88.0%) participants from the comic with social proof condition.

## 4.2 Bayesian Formulation

We use a Bayesian formulation of the problem of identifying suitable predictors for the messages in comic form. Kay et al. [36] provide a nice introduction on the appropriateness of Bayesian analysis for the HCI community. Bayesian analysis is attractive in our experiment due to two advantages: shifting the conversation from “did it work” to “how strong is the effect”; and benefits to small  $n$  studies.

Our experiment has three experimental groups: text, comic and comic with the social proof. One way to use classical statistical analysis for multiple groups would be to use ANOVA to see if

<sup>3</sup>From 2017 US Census Data <https://www.census.gov/quickfacts/fact/table/US>

the treatment (the use of comic) has any effect. Bayesian inference can sometimes differ from the standard ANOVA test widely used to compare treatment effects.

This is because ANOVA assumes equal variances within groups, and that the response within each group is drawn from a Normal distribution—in practice, both assumptions may be violated. It is straightforward in Bayesian Analysis to relax both assumptions: equal variances and Normality. The Normality assumption may be one reason why ANOVA and the  $t$ -test may be less sensitive to differences than Bayesian analysis [37, p. 470].

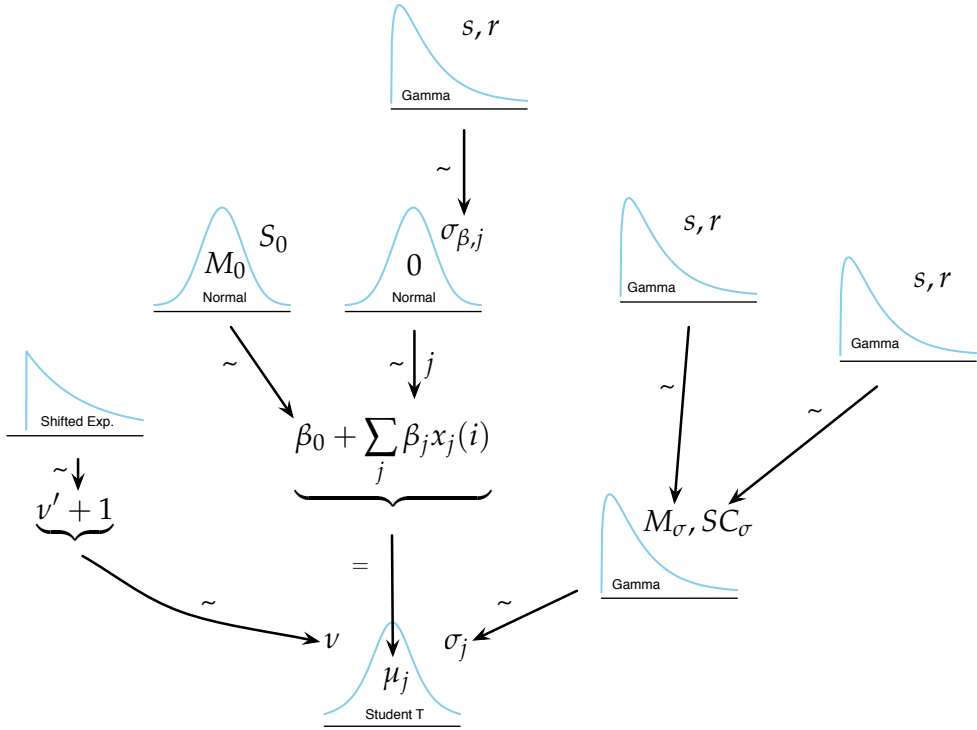


Fig. 5. Our proposed Hierarchical Bayesian model. The bottom of the figure shows the Student-t likelihood function for each experimental condition  $j$ . The likelihood function has three parameters:  $\nu$  (degrees of freedom),  $\mu_j$  (the mean for condition  $j$ ),  $\sigma_j$  (the scale for condition  $j$ ). The prior distribution for  $\nu$  is a shifted exponential distribution, so that  $\nu \geq 1$ . The prior distribution for the class mean  $\mu_j$  is a sum of Normal variables, where  $\beta_0$  is the central value across all three conditions, and  $\beta_j$  is the deflection corresponding to experimental condition  $j$ . Notice that  $x_j(i) = 1 \iff$  participant  $i$  is in condition  $j$ . We draw the variance of  $\beta_j$  from a Gamma distribution to ensure that  $\sigma_{\beta,j} > 0$ . We draw the scale parameters  $\sigma_j$  from the same Gamma distribution. We draw the mode ( $M_\sigma$ ) and the scale ( $S_\sigma$ ) parameters from a Gamma distribution to ensure  $M_\sigma > 0$ , and  $S_\sigma > 0$ . We set constants  $M_0, S_0, s, r$  to result in uninformative (i.e. conservative) priors.

Now, we discuss our Bayesian formulation (see Figure 5 for a graphical view of the model). There is one outcome variable  $y_{i|j}$ , the amount of donation to the charity by each participant  $i$ , under experimental condition  $j$ : text, comic, and comic with social proof.

In a Bayesian model formulation, we first need to define a likelihood function, to describe the data. We adopt the view advanced by McElreath [47] that a model represents an *epistemological*

claim (how the modeler views the data), not an *ontological* claim (i.e. a physical assumption about the world). Typically, the likelihood function is parametric, and we may treat each parameter as a random variable drawn from a distribution with parameters (this is the prior of the model parameter); thus Bayesian models can be hierarchical. In Bayesian modeling, the goal is to be conservative, and we express priors that are weakly informative—a technical term that informally means that the prior allows all possible values for the parameter, but does so in a way that allows for rapid model convergence.

We use a Student-t distribution to characterize the donations, since Figure 4 shows significant contributions for amount \$0 and \$5. A Student-t distribution is heavy-tailed, in the sense that the Student-t distribution doesn't fall off as quickly as does a Normal distribution. If we choose a Normal distribution to model contributions in each condition, then we would find it harder to explain contributions at \$0 and at \$5, without increasing the Normal distribution variance parameter. The Student-t distribution has three parameters: the degrees of freedom ( $\nu$ ), the experimental condition dependent mean ( $\mu_j$ ) and scale ( $\sigma_j$ ). Since our goal is to understand the average tendency to give to charity under the different conditions, and not to make predictions (as might be the case if we were trying to model donation with age as a predictor), the fact that the  $t$ -distribution is not bounded is less relevant here. Each of these parameters are random variables, and we need to define likelihood functions of each of them. Our Bayesian model (also see Figure 5):

$$y_{i|j} \sim \text{Student} - t(\nu, \mu_j, \sigma_j), \quad \text{likelihood function to model donation} \quad (1)$$

$$\nu \sim 1 + \exp(\lambda), \quad \text{degrees of freedom} \quad (2)$$

$$\mu_j \sim \beta_0 + \sum_j \beta_j x_j(i), \quad \text{modal contribution in each condition } j \quad (3)$$

$$\sigma_j \sim \Gamma(M_\sigma, SC_\sigma), \quad \text{scale parameter for condition } j \quad (4)$$

$$M_\sigma \sim \Gamma(s, r), \quad (5)$$

$$SC_\sigma \sim \Gamma(s, r), \quad (6)$$

$$\beta_0 \sim N(M_0, SD_0), \quad \text{average contribution across conditions} \quad (7)$$

$$\beta_j \sim N(0, \sigma_{\beta,j}), \quad \text{deflection from average contribution for condition } j \quad (8)$$

$$\sigma_{\beta,j} \sim \Gamma(s, r) \quad (9)$$

Equation (1) says that the response  $y_{i|j}$  of each group  $j$  is modeled as a Student –  $t$  distribution with mode  $\mu_j$ , scale  $\sigma_j$  and with  $\nu$  degrees of freedom; notice that this is a *drawing distribution*, not the  $t$ -test. The Student-t allows us to model a non-Normal outcome with a heavy tail; assuming  $y_{i|j}$  to be Normally distributed is equivalent to setting  $\nu = \infty$ . Now we explain the the main parameters of the model.

**Degrees of Freedom:** We draw the degrees of freedom  $\nu$  from a shifted exponential distribution, to ensure  $\nu \geq 1$

**Modal contribution  $\mu_j$  in each condition  $j$ :** The mode  $\mu_j$  corresponding to each group is drawn from a sum of Normally distributed random variables with  $\mu_j = \beta_0 + \sum_j \beta_j x_j(i)$ , where  $\beta_0$  corresponds to the overall mean contribution across experimental conditions, and where the group indicator variable  $x_j(i) = 1 \iff$  subject  $i$  belongs to group  $j$ . We model the overall group response  $\beta_0$  as a Normal distribution with mean  $\mu_0$  and variance  $\sigma_0$ . For each of the  $\beta_j$  for each group  $j$  in Equation (3),  $\beta_j$  is Normally distributed with mean  $\mu = 0$  and  $\sigma_{\beta,j}$  drawn from a Gamma distribution  $\Gamma(s, r)$  with shape parameter  $s$  and rate parameter  $r$ , ensuring  $\sigma_{\beta,j} > 0$ . The random variables  $\beta_j$  are centered around  $\mu = 0$  so that the group responses

are modeled as deflections around the overall mean  $\beta_0$ . Notice that we draw the variances  $\sigma_{\beta,j}$  of each group  $j$  independently from a Gamma distribution with same parameters  $(s, r)$ , implying that the variances (equivalently, the extent of deflections from the mean) for each predictor  $\beta_j$  can be different. The main advantage of using a Gamma distribution is that we can specify a non-zero mode, important in controlling shrinkage in hierarchical models.

**Scale  $\sigma_j$  of each condition  $j$ :** The scale  $\sigma_j$  of the likelihood function is drawn from a Gamma distribution  $\Gamma(M_\sigma, SC_\sigma)$ , with mode  $M_\sigma$  and scale  $SC_\sigma$ ; this prior on  $\sigma_j$  ensures that  $\sigma_j > 0$ . The mode  $M_\sigma$  and scale  $SC_\sigma$  are each drawn from two independent Gamma distributions  $\Gamma(s, r)$  with shape parameter  $s$  and rate parameter  $r$ . The advantage of this hierarchical approach is that the values of each element of  $\sigma_j$  inform the other elements. The “information sharing” among variables common to hierarchical Bayesian models and is an important reason why Bayesian models work so well with small datasets<sup>4</sup>.

**Constants:** The constants  $M_0, S_0, s, r$  are set so that the priors are generous but weakly informative so that despite exploring all possible values, we ensure rapid MCMC convergence.

Having introduced our model, now we examine the results of the data analysis.

### 4.3 Analysis

We analyzed the data using PyMC3 [61], a popular framework for Bayesian inference. Computational techniques for Bayesian inference use a stochastic sampling technique called Markov Chain Monte Carlo (MCMC) that samples the posterior distribution  $P(\theta|D)$ , where we want to estimate the parameters  $\theta$  given the observations  $D$ . In particular, we used the No-U Turn Sampler (NUTS) sampler.

Our analysis shows that the abstract comic form has a clear treatment effect over the corresponding text message. Figure 7 shows the contrast and the effect size of the contract for four cases. The first three columns of Figure 7 show the contrasts (top row) and the corresponding effect size (bottom row) between a pair of experimental conditions. The fourth column is a comparison between the text condition and the comic condition, by pooling together responses of the two comic conditions.

Let us examine one result (‘text’ v. ‘comic’) in detail; the analysis for the other cases follows a similar logic. The top left sub-figure shows that the contrast has a mode of 0.75 with the High Posterior Density (HPD<sup>5</sup>) interval of  $[0.20, 1.31]$ . That is, subjects most frequently donate \$0.75 more on in the comic condition than in the text condition, with 95% of the *increase in donations* lying between  $[\$0.20, \$1.31]$ . Since the HPD lies outside a significant ROPE (Region of Practical Equivalence)<sup>6</sup> of  $0 \pm 0.1$ , the result implies that there is a clear effect due to the treatment. Furthermore, the lower left sub-figure shows a moderate to medium effect size of 0.44; we consider an effect of 0.2 to be small, and an effect of 0.5 is a medium-sized effect. Since the distribution excludes

<sup>4</sup>The sharing of information causes the scale of each group  $\sigma_j$  to move towards the group variance, a phenomenon known as “shrinkage.”

<sup>5</sup>The HPD interval is the location of 95% of the posterior density. This is similar to, but different from the idea of the confidence interval used in non-Bayesian Statistics. In non-Bayesian Statistics, a confidence interval of say 95% is informally interpreted as “with 95% probability the parameter of interest lies in a specific interval; the tails are of equal width (i.e. 2.5%)”; the HPD on the other hand, is the *densest* interval covering 95% of the posterior. The HPD is guaranteed to include the most likely value, but this is not always true for confidence intervals; see McElreath [47, p. 57] for a simple example. For a more careful definition of the confidence interval, see Hoekstra et al. [29], which also includes a discussion on the misinterpretation of confidence intervals.

<sup>6</sup>Unlike non-Bayesian Statistics, where one can ask for example, if the two means for two treatments are different  $P(\mu_1 \neq \mu_2)$ , in Bayesian statistics, one asks if the HPD interval of the distribution  $P(\mu_1 - \mu_2)$ , that is, the distribution of the difference of the means of the two treatments, excludes an interval where we can consider the two treatments equivalent. This equivalence interval is domain dependent.



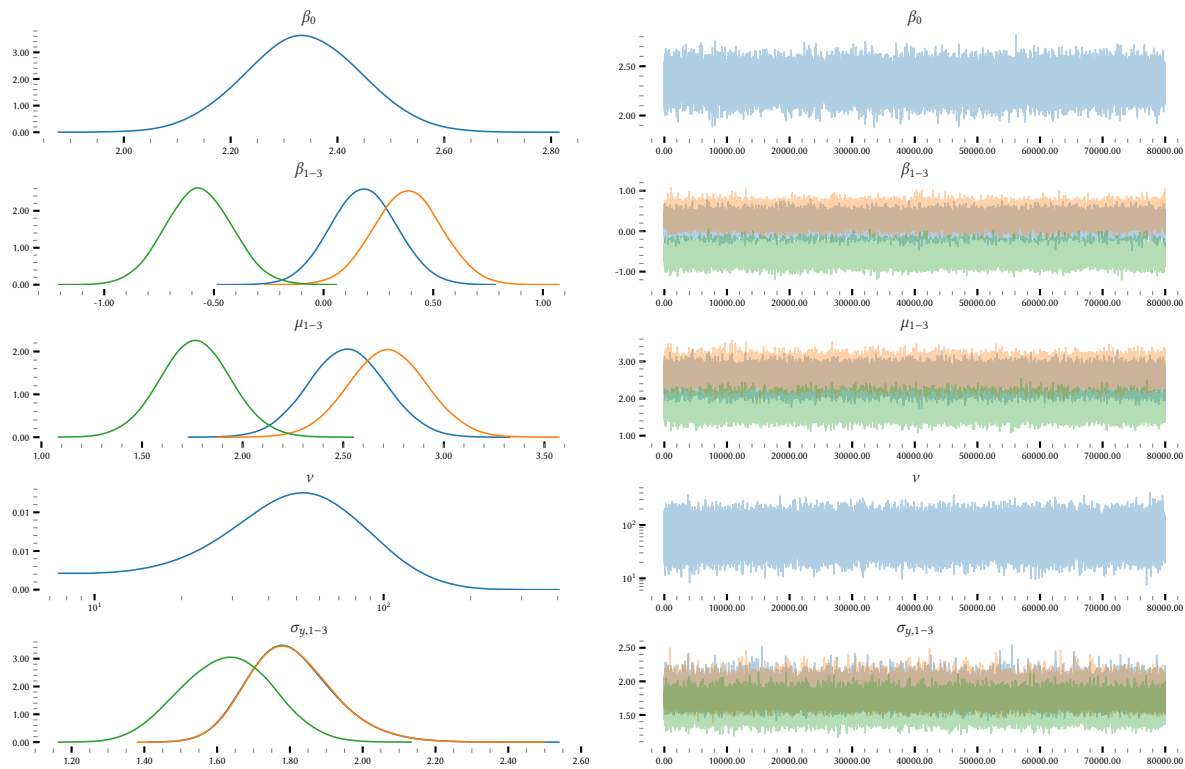


Fig. 6. Traceplot showing the results of the MCMC estimation. Left column shows the posterior distributions for  $\beta_0, \beta_{[1-3]}, \mu_{[1-3]}, \nu, \sigma_{[1-3]}$ , while the right column shows the corresponding traces. The posterior plots for  $\nu$ , the parameter corresponding to degrees of freedom has a modal value of  $\nu = 53$ . While the ideal case of  $\nu = \infty$  makes the Student-t distribution equivalent to the Normal distribution, in practice  $\nu \geq 30$  is used to test for Normality. Thus the response distributions  $y_{i|j}$  tend towards Normality. Notice that the mean of  $\sigma_2$  (green curve, fourth panel, left column), variance of the group response in the case of text, is less than  $\sigma_1$  or  $\sigma_3$  (variances of the group responses for the comic, comic+social norm cases respectively), justifying our modeling assumptions of unequal variances across groups; the variances for the two comic conditions are nearly identical. The Gelman-Rubin statistic  $\hat{R}$  was around 1, indicating that the different sampling chains converged. Furthermore, the effective sample size of all parameters was greater than 10,000.

the ROPE interval  $[0 \pm 0.1]$ , half of the small-sized effect value of 0.2, the discovered effect size is statistically significant.

The second column contrasts the comic with social proof case ('comic+social') with the plain text ('text'). We find that the contrast has a mode of 0.95 with the High Posterior Density (HPD) interval of  $[0.38, 1.52]$ . The modal effect size is 0.55, with an HPD interval of  $[0.21, 0.89]$  which excludes a ROPE of  $0 \pm 0.1$  indicating a significant, slightly larger than a medium-sized effect.

The third column contrasts the donations in the comic ('comic') condition against the comic with the social proof ('comic+social'). While the contrast is positive with a mode of 0.19, notice that the HPD interval  $[-0.37, 0.77]$  overlaps a ROPE of  $0 \pm 0.1$  implying that the observed differences are not significant. The corresponding modal effect size of this contrast is 0.11 with an HPD of  $[-0.20, 0.43]$  with about 21.9% of the HPD lying to the left of 0, implying that the observed effect size is not significant.

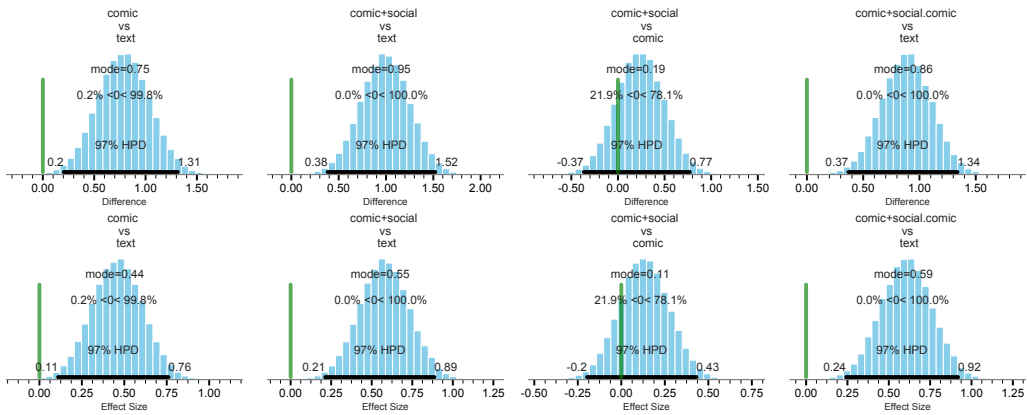


Fig. 7. The figure shows contrasts between the three experimental conditions ('text', 'comic', 'comic+social'). We also show in the fourth column contrasts between the text condition and pooled case of the two comic conditions to highlight the effect of using the abstract comic. Since we are highlighting contrasts, each sub-figure also shows a green line located at 0. The main finding: the comic treatment creates a medium to large sized effect (mode=0.59, pooled case, fourth column, bottom) when compared to the 'text' condition. There is a slight increase in effect size when using the comic with social proof ('comic+social', second column, bottom) in comparison to the case when there is a comic without the social proof ('comic', first column, bottom). However, these differences are not significant, since the about 21.9% of the HPD of the contrast between the 'comic' and the 'comic+social' case lies to the left of 0.0 (second column).

The fourth column of Figure 7 compares the text condition ('text') with the comic condition by pooling the responses in the two comic conditions. The pooling of the 'comic' and the 'comic+social' data is straightforward in Bayesian analysis. Since the MCMC *jointly* estimates the posteriors of all variables, we can create a new variable that averages the values of the posteriors of the 'comic' and 'comic+social' cases for every step of the MCMC, to represent an averaged 'comic' condition. We find that the contrast has a mode of 0.86 with the High Posterior Density (HPD) interval of  $[0.37, 1.34]$ . Notice that the HPD interval is narrower than the corresponding HPD intervals comparing ('text' v. 'comic') and ('text' v. 'comic+social') The modal effect size is 0.59, with an HPD interval of  $[0.24, 0.92]$  which excludes a ROPE of  $0 \pm 0.1$  indicating a significant, a medium to large sized effect, where the large effect size is 0.8.

In this section, we presented a Bayesian model to analyze the overall effect of using an abstract-comic to persuade people making charitable donation decisions, and understand the effect of

adopting persuasive techniques in the abstract comic form. The results show that the use of abstract-comic produces a significant, medium to large modal effect (0.59) in persuading participants to donate. Although abstract-comic with social proof ('comic+social') can produce a larger effect than the case without social proof ('comic'), the contrast between abstract-comic and abstract-comic with social proof is not significant. The code and data are available online <sup>7</sup>.

## 5 MODEL CRITICISM

In this section, we begin with a short section explaining our choice to use Bayesian analysis. Then in 5.2, we examine model convergence and Normality.

### 5.1 Why Bayesian?

In a recent paper, Kay et al. [36], make a persuasive argument that Bayesian methods are better suited to the HCI community, including making the case that Bayesian methods allow for replicating the results and improving the strength of the conclusions by using previous outcomes as priors. We would add two reasons, in addition to those by Kay et al. [36] to explain our decision to use Bayesian models.

**Transparency:** With a Bayesian model, the researcher foregrounds all the aspects of the model; there are no modeling assumptions that need checking, not already foregrounded in the model description. Non-Bayesian statistics are powerful tools, and when used by an experienced statistician, they can dramatically reduce the process of inference. However, for researchers who wish to investigate their findings without access to a statistician, they need to be careful of the assumptions of the different tests: Normality ( $t$ -test); heteroscedasticity (e.g., ANOVA) and ensuring that the data satisfy the assumptions. Omitting the right sequence of analysis can lead to inferences not supported by the data.

**Small  $n$  studies:** A Bayesian model is valid at every value of  $n$ ; we do not have to wait for  $n \geq 30$  to satisfy assumptions of say Normality. For small  $n$  values, the result is of course affected by the choice of the prior; but by using weakly informative priors, we can ensure that the prior doesn't dominate inference. Furthermore, when Bayesian models use maximum entropy likelihood functions (e.g., members of the exponential family, that include the Normal distribution and the gamma distribution), we make the *most conservative* inference given the data. See McElreath [47, Chapter 9] for an excellent description of the use of maximum entropy models in Bayesian analysis.

### 5.2 Model Convergence and Normality

In this section, we examine model convergence and our assumptions about Normality.

The model shows good convergence, as evidenced by the traceplot in Figure 6. The Gelman-Rubin statistic  $\hat{R}$  was around 1, indicating that the different sampling chains converged. Furthermore, the effective sample size of all parameters was greater than 10,000.

We used the  $t$ -distribution to model likelihood motivated by the high contributions at \$0 and \$5 in the data (c.f. Figure 4) implying that a heavy-tailed distribution may be a better likelihood function than a Normal distribution. Let us examine the posterior distribution for  $\nu$ , the degrees of freedom of the  $t$ -distribution. As a reminder, the  $t$ -distribution is equivalent to the Normal distribution when  $\nu = \infty$ . We can see that while the 95% HPD lies between [17.37, 147.19], less than 7% of the posterior lies below  $\nu = 30$ , the traditional rule-of-thumb in non-Bayesian statistics for use of the Normal distribution. Since there is only a 7% chance that the degrees of freedom  $\nu \leq 30$ ,

<sup>7</sup><https://github.com/CrowdDynamicsLab/Abstract-Comic-Persuasion>

instead of using the Student- $t$  likelihood function, we could use a Normally distributed likelihood function with unequal variance.

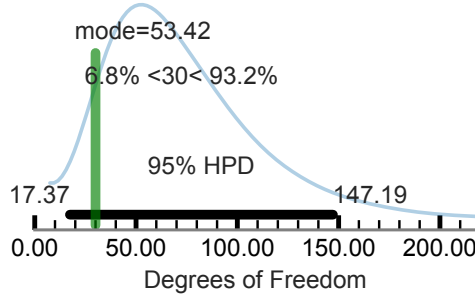


Fig. 8. The posterior distribution for  $\nu$ , the normality parameter of the  $t$ -distribution. When  $\nu = \infty$  the  $t$ -distribution is identical to the Normal distribution. The posterior distribution shows a vertical green bar for  $\nu = 30$ , the traditional cut-off condition on degrees of freedom, for using a Normal distribution. The mode  $\nu = 53.42$ , and the posterior probability distribution implies that only about 7% of the posterior lies to the left of the green line (i.e.  $P(\nu \leq 30) \approx 0.07$ ), implying that the assuming that the likelihood function to be Normally distributed will give similar inference.

Our posterior predictive check shows that the  $t$ -distribution models well the heavy tails, and shows that the variances for the comic conditions are different from the text condition. Furthermore, the comparison between the hierarchical model with unconstrained variances and the model with equal variances shows that the hierarchical model is better at explaining the observations. We discuss posterior prediction checks (Section 8.1) and alternative models (Section 8.2) in detail in the Appendix.

To summarize, we discussed transparency and utility in small- $n$  studies as motivation for our use of Bayesian modeling. The model shows good convergence with the Gelman-Rubin statistic  $\hat{R}$  was virtually identical to 1.0 for all parameters; the effective sample size was greater than 10,000 for all parameters. The modal value of  $\nu$ , the degrees of freedom parameter was around 53, suggesting that we could also use the Normal likelihood function.

## 6 DISCUSSION

We will first summarize our findings and discuss experimental variations. Then, we will propose a framework for algorithmically synthesizing persuasive messages into the abstract comic form, discuss design implications, and identify limitations.

### 6.1 Experiments: Findings and Alternative Contexts

From asking individuals to act to appealing for charitable donations, text messages have been widely used in simulating behaviors. Scholars from psychology have shown how variations in the text messages construction alter decisions. In this study, we explored and examined the role of the abstract comic form, a highly expressive, affective medium in communicating persuasive messages. To test the effectiveness of abstract comic persuasive messages, we persuaded individuals to make online charitable donations, a common public good dilemma. In public good dilemma, due to the non-exclusive and non-rivalrous nature of public goods, persuading individuals to contribute is hard and crucial. Also, online charitable donation tasks not only avoid confounding factors such

as habit formation which exists in other persuasion tasks (e.g., exercise and healthy diet) but also assure the ecological validity of our study as charity and organizations often solicit donations online. In our study, we compared the persuasive power of text messages, comic messages, and comic messages with social proof in asking charitable donations to public health research (e.g., the Organization for Autism Research).

**Comic vs. Plain Text:** Our results show that study participants prefer persuasive messages in abstract comic form over plain text. When making the charitable donation between \$0 to \$5, study participants donated \$0.86 more if they read the persuasive message in an abstract comic form (see Figure 7, fourth column). The results demonstrated the persuasive power of abstract comic in stimulating behaviors in pro-social decisions. Our findings are consistent with prior research on visual persuasive messages and the benefits of comics in communication. One potential explanation is that study participants were more attracted by the comic strip and projected themselves onto the character. When the projection happened, the persuadee may be able to digest the information better which stimulated them to donate more to the Organization for Autism Research.

**Social Proof Condition:** When comparing between the comic condition and comic with the social proof condition, although study participants who read the comic with social proof donated more on average (\$0.19 more) than the comic only condition, our analysis results do not indicate a significant difference between the two conditions (see Figure 7, third column). There are two potential explanations. First, it is possible that the design of our comic strip did not strongly signify the idea of social proof — we simply introduced the statistic from our pilot study in the text bubble. Other comic designs, e.g., adding other donators as comic characters, may make the social proof more salient. Another potential explanation is that the influence from other study participants is not strong enough. As Goldstein et al. [27] discovered in the hotel towel re-use study, the stronger the relationship between the subject and the reference group of individuals is, the more persuasive the message is. The relationship between people on Amazon Mechanical Turk are likely to be weaker than the hotel guests in Goldstein et al. [27]’s towel re-use study.

We conjecture that using a person’s social network as part of the social proof (after seeking consent, for example, to use the person’s Facebook profile information) may increase the contribution, but needs further research. We know from Cialdini’s seminal work on persuasion that the knowledge of friends performing similar behavior is persuasive. We also know from Small et al. [67]’s work on charitable donation that named entities are more powerful in persuasion than abstract statistics (e.g. “Jane Doe is a hurricane survivor; please help” vs. “3000 people homeless due to the hurricane; please help”). If the person has given consent, a charity could find out in real time, the Facebook friends of the person who have loaded the web-page, and compute the fraction who have donated and use that number in the social proof. Or a charity can use tie-strength [26], if available, and name the closest friend who has donated to the charity in the social proof message. However, it may backfire as we know from recent work [23] that when advertisers display information that subjects find “creepy” (as demonstration of knowledge by the charity of the subject’s closest friend, or even statistics of the person’s Facebook friends circle contributions), the subjects may find the charity less trustworthy and may donate less.

**Donation Amount:** Another natural question to ask is whether the maximum donation amount of \$5 in our study has any influence on our findings (e.g., will people make the same donation decision if they can donate more?). Prior studies in behavioral economics suggested that small-stake experiment in developed countries can be replicated in developing countries where the stake becomes a significant portion of participants’ weekly income [8, 9, 34]. Additionally, Post et al. [58] showed that the results from small-stake experiments would hold with higher stakes

in developed countries as well. Those results seem to suggest that our findings will hold with higher donation amount. However, the best way to confirm is through actual experiments.

**Other Media Forms:** While a direct comparison between the abstract comic form and other visual forms such as image and video is beyond the scope of the paper, we discuss two advantages of using the abstract comic to persuade. There are two points to consider: the ease of making a connection with a viewer (the abstract comic form vs. images or video); the infrastructure to create tailored messages. First, as McCloud [44] points out, when using the abstract comic form, the reader projects themselves onto the character whereas a concrete and detailed representation of the person in the comic can cause the reader to develop a sense of the “other” - that is, they are viewing someone else enacting a scene. Images and videos often use actual human models which will create a sense of the “other” for the viewer. So, images/videos need to connect with the viewer through other demographic, cultural and local cues. For example, in advertising, to allow recipients of ads to identify or resonate with the message context, advertisers try to ensure that the models who appear in images match the target demographic. In contrast, when viewing an abstract comic, readers fill in the gaps with their own characteristics, and they see themselves as the protagonist. Second, our infrastructure (discussed next in Section 6.2) makes it straightforward for a charity to compose a multi-panel abstract comic. The infrastructure to create tailored imagery and videos may be expensive for the charity. Today, advertisers tailor the imagery (usually the individual in the image) in online banner ads (embedded on a web page) to the demographics of the person viewing the ad. Matching the demographics requires that advertisers access the information stored in a browser cookie against third-party information about the person in real-time. This real-time switching of imagery is expensive: it requires a manually curated image collection corresponding to individuals matching to different demographic segments of interest. And, to personalize the message (e.g. with the social proof), they would need to have the infrastructure to compose the message on the fly. Furthermore, algorithmic synthesis of persuasive, personalized videos is more challenging. One possibility that we do consider as part of future work: animations that involve the abstract comic form.

**Experiment Flow in the Real World:** We consider three online charitable donation scenarios related to how the charitable organization knows the person asked to make a donation. In each case, consistent with prior work [57], the charity needs to establish context.

First, the person may be familiar with the charity through their lived experience. For example, in the case of Wikipedia, the person may use the site frequently. A person who regularly uses the Wikipedia site is unlikely to need a video describing Wikipedia benefits. In such cases, we envision using the three-panel comic instead of the desktop banner ad that Wikipedia uses at the moment.

Second, the person may be a member of a charity or a non-profit organization (e.g. museums). For non-profit organizations with memberships, they can recognize a member has logged onto the organization website. In this case, detailed context may be less necessary and organizations could use the three-panel comic in their regular donation page, along with social proof describing the fractions of other members who have contributed.

Third, the person may not know the charity. For charities who need to ask for donations from people without neither lived experience nor memberships, they need to establish context from scratch. They may establish context through videos, a page with text and images or even a physical mailer. After watching the videos or reading the page, they may show the person a comic message. For the physical mailer, the charity would include its usual textual description and imagery followed by the abstract three-panel comic. After the comic panel, they may direct the person opening the mailer make a contribution on a web-page.



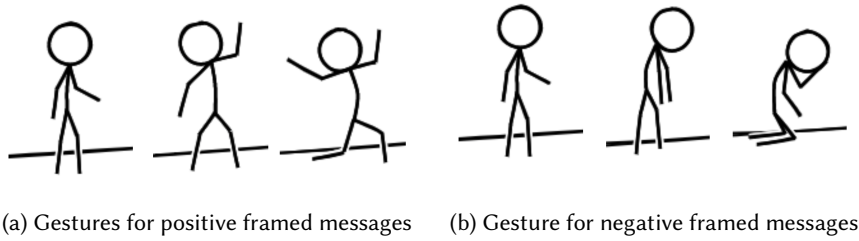


Fig. 9. Different character gestures to communicate various levels of emotional intensity. The left figure shows gestures from neutral to the happiest. The right figure shows gestures from neutral to the most frustrated.

For the first two scenarios, a charitable organization (e.g. Wikipedia) can determine regular use of its website or determine if the person visiting the web-page is an organization member through the use of advertising cookies with permission.

## 6.2 Framework for Algorithmically Synthesized Abstract-Comic Persuasive Messages

Our study showed the persuasive power of abstract comics in encouraging people to make pro-social decisions. However, one drawback of using visual stimulus in persuasion is the cost during creation processes. Although compared to creating other persuasive visual stimuli such as videos or complex graphical illustration, due to its simplicity, abstract comics requires much less effort. In this section, we propose a framework that allows full/semi- automatic generation of abstract persuasive messages and identify crucial features that need to be addressed in future work. We believe such a tool will lower the barrier for the persuader to take advantage of the abstract comic (as demonstrated in our study) in encouraging individuals to act in public good dilemmas.

In our study, we created a comic generator with existing packages including “cmx.io” [28] and “rough.js” [66] to generate three-panel abstract comic strips. With several pre-defined character gestures (see Figure 9), the generator only requests the text input from the persuader to create the comic message.

We believe the comic generator built in this study can be further developed as a framework for algorithmic synthesis. Now, we identify crucial features that need to be addressed in future research. First, the framework should be able to automatically select the character gesture that best fits the persuasive message’s context. For example, when the message receiver was told good news, his/her gesture should reflect that conflict. The appropriate mapping will create a natural and coherent comic message which is the key for an expressive message. We need future research to realize such a method. Second, the framework should be able to use other comic elements, such as inter-character distance and shading, to create persuasive comics. To achieve this feature, we need to understand how different elements affect the persuasiveness of the abstract comic messages. Third, the framework should be able to personalize the persuasive comics by using persuadee’s behavioral data to inform comic contents. We could derive the statistics for the social proof from the behaviors of a person’s friends and the data from his/her own activities.

## 6.3 Design Implications

Our main design implications are two-fold. First, when encouraging individuals to contribute to public goods, non-profit organizations and governmental agencies ought to consider using abstract comic in their online messaging campaign to persuade. From our results, using abstract comics can,

in particular, persuade people to act in public-goods dilemmas that require single-shot decisions (online charitable donations). Additionally, when constructing persuasive comics, it is worthwhile to consider to incorporate social proofs. In the study, the purpose of the video is to inform participants about the context. Charities could use the comic after they've established context as discussed in Section 6.1.

Second, our results also inform the development of a computational framework that can algorithmically synthesis textual persuasive messages into abstract comic form and use data-driven methods to personalize the message to further increase messages' persuasive power.

#### 6.4 Limitations & Future Work

**Distant and Non-exclusive Task:** Although our study asked participants to make the decision with a real cost, the decision domain is limited to only one scenario, online charitable giving. In this task, the participant's reward is distant and non-exclusive (e.g., participant's donation decision won't bring any immediate reward and won't exclude the participant from the research outcome). Although both characteristics help us clearly test the persuasive benefits of abstract comics, they limited our generalizability. We need future research to understand the pros and cons of using abstract comic messages in persuasive tasks with different reward characteristics. For example, for persuasive goals in the quantified-self movement [22, 15] such as exercise and dieting, the reward is distant (e.g., people's health won't be improved immediately after exercise) but exclusive (e.g., healthy life situation mainly benefit the individual him/herself). Moreover, our persuasion task is for an online charitable donation. Future work is needed to extend our result to offline charitable donation solicitation where persuasive text messages, such as mailers, also often used.

**Study on Amazon Mechanical Turk:** Consistent with prior research [39, 62, 68, 3, 10] that used Amazon Mechanical Turk to study charitable donations, we recruited our participants from Amazon Mechanical Turk.

Although research shows that populations on Amazon Mechanical Turk are diverse and can mirror the US population [11, 6, 7], there are concerns about Amazon Mechanical Turk's sample representativeness [38, 56]. One potential solution is to use panel population where the panel company claimed to offer a more diverse and representative sample. However, this method also has concerns in that the researcher can not directly cross-validate the sample's representativeness. Another solution is using stratified sampling, but it is challenging to classify every member of the population into a subgroup properly.

Another limitation from the use of Amazon Mechanical Turk subjects in our study is that prior research [56, 55, 35] indicates that compared to other recruiting methods, study participants from Amazon Mechanical Turk are more sensitive to monetary rewards as the monetary reward is their prime motive for participation. Since we ask our Amazon Mechanical Turk study participants to donate from their own prospective rewards, we may see a lower average donation across all conditions than asking the population sample less sensitive to monetary rewards. An alternative sampling method we could consider in the future work is recruiting participants from our own social network. We conjecture that we may see higher average contributions in each of the three conditions, since participants on social network may be less sensitive to monetary rewards than those on Amazon Mechanical Turk. However, we need to carefully control for other potential confounding factors such as study participant's knowledge about our experiment goal.

**Comic Message Construction:** In our study, we constructed the comic messages in the XKCD comic style. Although the XKCD style is easily recognizable, there are many different ways to create abstract comics. Comic, as a creative art form, has rich semantics and vocabularies

to communicate ideas [45]. Out of necessity, we limited the complexity our comics strip (e.g., number of characters, gestures, and the number of panels). Future research is needed to explore and understand how other comic elements, for example, inter-character distance and character's gesture, affect the persuasiveness of the comic. Furthermore, while simplicity grants us the possibility to automate the generation process, future technology may allow us to generate more complicated persuasive abstract comics.

**Other Experimental Contexts:** What are other contexts to which our study results may apply? Although our study examined charitable donation decisions, a single shot task with distant, non-exclusive rewards, certain health related tasks may be candidates to study next. Recent work by Bushar et al. [13] on using SMS text messages to persuade pregnant mothers to take influenza vaccinations suggests that those who received the text messages were more likely to report taking a vaccination than those who did not receive a message. Further research will be merited to determine if for this single-shot task, with exclusive, non-immediate reward, an abstract comic (e.g. embedded as a link in the SMS) succeeds in increasing vaccination rates for pregnant women over the plain text.

## 7 CONCLUSION

Inspired by a rich history in persuasive text messages construction and benefits of abstract comics in communication, this paper examined if the abstract comic form is more persuasive than the corresponding plain text in the context of encouraging participants to donate for charitable causes. We conducted a field study on Amazon Mechanical Turk with 307 participants. In the study, participants received one of the three persuasive messages designed to ask for a donation to the Organization of Autism Research, a persuasive text message, a three-panel comic strip, and a three-panel comic strip incorporating the idea of social proof, and made donation decisions with real money. We analyzed the results using a hierarchical Bayesian framework that allows for understanding effect sizes, as well as transparent and helpful in small- $n$  studies. The results shows convincingly that the three-panel abstract comic is more persuasive than the text (a medium to large effect size = 0.59). We also show that while the comic with social proof increases the donation level over the comic without the norm, the effect size is very small (0.11) and the increase is not significant. To summarize, the comic form significantly increases donations over the plain text, but the presence of the norm is not effective. We caution that the result holds for single-shot, public goods tasks; the value of the social proof in the comic, for exclusive tasks with distant rewards such as exercise, or dieting needs future research. The main implication of our work is that non-profits and Governmental agencies ought to consider using abstract comic in their online campaign as they work to alleviate public goods dilemmas. We believe that these agencies can easily include the use of the comic form as part of their overall messaging strategy because the simplicity of the abstract comic form allows it to be easily synthesized and to additionally incorporate social proofs.

As next steps, we plan to develop an algorithmic framework that automatically maps a person's behavioral data (e.g. amount walked this week) to a three-panel persuasive comic. We also plan to conduct longitudinal field experiments with an emphasis on storytelling where individuals receive three-panel comics over time, and comics are connected with a storyline.

## ACKNOWLEDGMENTS

We thank our study participants who joined our study. We acknowledge reviewers for their insightful feedback.

## REFERENCES

- [1] Naveen Ambalee and Tung Bui. 2011. Harnessing the influence of social proof in online shopping: the effect of electronic word of mouth on sales of digital microproducts. *International Journal of Electronic Commerce*, 16, 2, 91–114.
- [2] James Andreoni. 1990. Impure altruism and donations to public goods: a theory of warm-glow giving. *The economic journal*, 100, 401, 464–477.
- [3] Antonio A Arechar, Gordon T Kraft-Todd, and David G Rand. 2017. Turking overtime: how participant characteristics and behavior vary over time and day on amazon mechanical turk. *Journal of the Economic Science Association*, 3, 1, 1–11.
- [4] American Psychiatric Association et al. 2013. *Diagnostic and statistical manual of mental disorders (DSM-5)*. American Psychiatric Pub.
- [5] Gary S Becker. 1974. A theory of social interactions. *Journal of political economy*, 82, 6, 1063–1093.
- [6] Tara S Behrend, David J Sharek, Adam W Meade, and Eric N Wiebe. 2011. The viability of crowdsourcing for survey research. *Behavior research methods*, 43, 3, 800.
- [7] Adam J Berinsky, Gregory A Huber, and Gabriel S Lenz. 2012. Evaluating online labor markets for experimental research: amazon. com’s mechanical turk. *Political analysis*, 20, 3, 351–368.
- [8] Hans P Binswanger. 1980. Attitudes toward risk: experimental measurement in rural india. *American journal of agricultural economics*, 62, 3, 395–407.
- [9] Hans P Binswanger et al. 1981. Attitudes toward risk: theoretical implications of an experiment in rural india. *Economic Journal*, 91, 364, 867–890.
- [10] Pablo Braas-Garza, Valerio Capraro, and Ericka Rascon-Ramirez. 2018. Gender differences in altruism on mechanical turk: expectations and actual behaviour. *Economics Letters*, 170, 19–23.
- [11] Michael Buhrmester, Tracy Kwang, and Samuel D Gosling. 2011. Amazon’s mechanical turk: a new source of inexpensive, yet high-quality, data? *Perspectives on psychological science*, 6, 1, 3–5.
- [12] John J Burnett. 1981. Psychographic and demographic characteristics of blood donors. *Journal of Consumer Research*, 8, 1, 62–66.
- [13] Jessica A. Bushar, Juliette S. Kendrick, Helen Ding, Carla L. Black, and Stacie M. Greby. 2017. Text4baby influenza messaging and influenza vaccination among pregnant women. *American Journal of Preventive Medicine*, 53, 6, 845–853.
- [14] Stephen Cary. 2004. *Going graphic: Comics at work in the multilingual classroom*. Heinemann Portsmouth, NH.
- [15] Eun Kyoung Choe, Nicole B. Lee, Bongshin Lee, Wanda Pratt, and Julie A. Kientz. 2014. Understanding quantified-selfers’ practices in collecting and exploring personal data. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI ’14)*. ACM, Toronto, Ontario, Canada, 1143–1152.
- [16] Robert B Cialdini. 1993. Influence: the psychology of persuasion. *New York, Morrow*.
- [17] Robert B. Cialdini and Noah J. Goldstein. 2004. Social influence: compliance and conformity. *Annual Review of Psychology*, 55, 1, 591–621. PMID: 14744228.
- [18] Sunny Consolvo et al. 2008. Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1797–1806.
- [19] S Cotterill, P John, and L Richardson. 2010. The impact of a pledge campaign and the promise of publicity on charitable giving: a randomised controlled trial of a book donation campaign. In *Randomised Controlled Trials in the Social Sciences Conference, York, September 2010*.
- [20] Mette Trier Damgaard and Christina Gravert. 2017. Now or never! the effect of deadlines on charitable giving: evidence from two natural field experiments. *Journal of Behavioral and Experimental Economics*, 66, 78–87.
- [21] Sanorita Dey, Brittany Duff, Karrie Karahalios, and Wai-Tat Fu. 2017. The art and science of persuasion: not all crowdfunding campaign videos are the same. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 755–769.
- [22] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS ’14)*. ACM, Vancouver, BC, Canada, 667–676.
- [23] Motahhare Eslami, Sneha R Krishna Kumaran, Christian Sandvig, and Karrie Karahalios. 2018. Communicating algorithmic process in online behavioral advertising. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 432.
- [24] Daniel C Feiler, Leigh P Tost, and Adam M Grant. 2012. Mixed reasons, missed givings: the costs of blending egoistic and altruistic reasons in donation requests. *Journal of Experimental Social Psychology*, 48, 6, 1322–1328.
- [25] Andrew Gelman, Jessica Hwang, and Aki Vehtari. 2014. Understanding predictive information criteria for bayesian models. *Statistics and Computing*, 24, 6, (Nov. 2014), 997–1016.
- [26] Eric Gilbert and Karrie Karahalios. 2009. Predicting tie strength with social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’09)*. ACM, Boston, MA, USA, 211–220.

- [27] Noah J Goldstein, Robert B Cialdini, and Vidas Griskevicius. 2008. A room with a viewpoint: using social norms to motivate environmental conservation in hotels. *Journal of consumer Research*, 35, 3, 472–482.
- [28] Antonin Hildebrand. 2013. Comix i/o. (Feb. 2013).
- [29] Rink Hoekstra, Richard D Morey, Jeffrey N Rouder, and Eric-Jan Wagenmakers. 2014. Robust misinterpretation of confidence intervals. *Psychonomic bulletin & review*, 21, 5, 1157–1164.
- [30] Peter S Houts, Cecilia C Doak, Leonard G Doak, and Matthew J Loscalzo. 2006. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. *Patient education and counseling*, 61, 2, 173–190.
- [31] R Mark Isaac, Kenneth F McCue, and Charles R Plott. 1982. Public goods provision in an experimental environment.
- [32] Aarti Iyer and Julian Oldmeadow. 2006. Picture this: emotional and political responses to photographs of the kenneth bigley kidnapping. *European Journal of Social Psychology*, 36, 5, 635–647.
- [33] Marco A. Janssen, Allen Lee, and Hari Sundaram. 2016. Stimulating contributions to public goods through information feedback: some experimental results. *PLoS ONE*, 11, 7, (July 2016), 1–16.
- [34] Steven J Kachelmeier and Mohamed Shehata. 1992. Examining risk preferences under high monetary incentives: experimental evidence from the people's republic of china. *The American Economic Review*, 1120–1141.
- [35] Nicolas Kaufmann, Thimo Schulze, and Daniel Veit. 2011. More than fun and money. worker motivation in crowdsourcing-a study on mechanical turk. In *AMCIS number 2011*. Vol. 11. Detroit, Michigan, USA, 1–11.
- [36] Matthew Kay, Gregory L. Nelson, and Eric B. Hekler. 2016. Researcher-centered design of statistics: why bayesian statistics better fit the culture and incentives of hci. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, Santa Clara, California, USA, 4521–4532.
- [37] John Kruschke. 2014. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.
- [38] Richard N Landers and Tara S Behrend. 2015. An inconvenient truth: arbitrary distinctions between organizational, mechanical turk, and other convenience samples. *Industrial and Organizational Psychology*, 8, 2, 142–164.
- [39] Yu-Hao Lee and Gary Hsieh. 2013. Does slacktivism hurt activism?: the effects of moral balancing and consistency in online activism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 811–820.
- [40] Jennifer Mankoff et al. 2010. Stepgreen. org: increasing energy saving behaviors via social networks. In *Fourth International AAAI Conference on Weblogs and Social Media*.
- [41] Gerald Marwell and Ruth E Ames. 1981. Economists free ride, does anyone else?: experiments on the provision of public goods, iv. *Journal of public economics*, 15, 3, 295–310.
- [42] Mizuki Matsubara, Olivier Augereau, Charles Lima Sanches, and Koichi Kise. 2016. Emotional arousal estimation while reading comics based on physiological signal analysis. In *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding*. ACM, 7.
- [43] Nicole D. Mayer and Zakary L. Tormala. 2010. Think versus feel framing effects in persuasion. *Personality and Social Psychology Bulletin*, 36, 4, 443–454.
- [44] Scott McCloud. 2011. *Making comics*. Harper Collins.
- [45] Scott McCloud. 1993. *Understanding comics*. William Morrow Paperbacks.
- [46] Jason E. McDermott, Matthew Partridge, and Yana Bromberg. 2018. Ten simple rules for drawing scientific comics. *PLoS Computational Biology*, 14, 1.
- [47] Richard McElreath. 2015. *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.
- [48] Doug McKenzie-Mohr. 2011. *Fostering sustainable behavior: An introduction to community-based social marketing*. New society publishers.
- [49] Paul Messaris. 1997. *Visual persuasion: The role of images in advertising*. Sage.
- [50] Cees Midden, Teddy McCalley, Jaap Ham, and Ruud Zaalberg. 2008. Using persuasive technology to encourage sustainable behavior. *Sustainability WS Pervasive*, 113.
- [51] Edward Miguel and Mary Kay Gugerty. 2005. Ethnic diversity, social sanctions, and public goods in kenya. *Journal of public Economics*, 89, 11-12, 2325–2368.
- [52] Randall Munroe. 2009. Xkcd: a webcomic of romance, sarcasm, math, and language. *Creative Commons Attribution-Noncommercial*, 2.
- [53] Richard E Nisbett and Lee Ross. 1980. Human inference: strategies and shortcomings of social judgment.
- [54] Mancur Olson. 2009. *The logic of collective action*. Vol. 124. Harvard University Press.
- [55] Gabriele Paolacci and Jesse Chandler. 2014. Inside the turk: understanding mechanical turk as a participant pool. *Current Directions in Psychological Science*, 23, 3, 184–188.
- [56] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. 2010. Running experiments on amazon mechanical turk. *Judgment and Decision making*, 5, 5, 411–419.
- [57] Edgar A Pessemier, Albert C Bemmaor, and Dominique M Hanssens. 1977. Willingness to supply human body parts: some empirical results. *Journal of Consumer Research*, 4, 3, 131–140.

- [58] Thierry Post, Martijn J Van den Assem, Guido Baltussen, and Richard H Thaler. 2008. Deal or no deal? decision making under risk in a large-payoff game show. *American Economic Review*, 98, 1, 38–71.
- [59] Organization for Autism Research. 2018. Run for autism - david brown. (July 2018).
- [60] S Rumsey, DP Hurford, and AK Cole. 2003. Influence of knowledge and religiousness on attitudes toward organ donation. In *Transplantation proceedings* number 8. Vol. 35. Elsevier, 2845–2850.
- [61] John Salvatier, Thomas V. Wiecki, and Christopher Fonnesbeck. 2016. Probabilistic programming in python using pymc3. *PeerJ Computer Science*, 2, (Apr. 2016), e55.
- [62] Timothy J Saunders, Alex H Taylor, and Quentin D Atkinson. 2016. No evidence that a range of artificial monitoring cues influence online donations to charity in an mturk sample. *Royal Society open science*, 3, 10, 150710.
- [63] Bernd H Schmitt, Nader T Tavassoli, and Robert T Millard. 1993. Memory for print ads: understanding relations among brand name, copy, and picture. *Journal of Consumer Psychology*, 2, 1, 55–81.
- [64] P Wesley Schultz, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein, and Vidas Griskevicius. 2007. The constructive, destructive, and reconstructive power of social norms. *Psychological science*, 18, 5, 429–434.
- [65] Ted Selker, Shih-Yuan Yu, Che-Wei Liang, and Jane Hsu. 2015. Sweetbuildinggreeter: a demonstration of persuasive technology for public space. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 475–486.
- [66] Preet Shih. 2017. Rough.js. (Feb. 2017).
- [67] Deborah A Small, George Loewenstein, and Paul Slovic. 2007. Sympathy and callousness: the impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes*, 102, 2, 143–153.
- [68] Abigail B Sussman, Eesha Sharma, and Adam L Alter. 2015. Framing charitable donations as exceptional expenses increases giving. *Journal of Experimental Psychology: Applied*, 21, 2, 130.
- [69] Amos Tversky and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *Science*, 211, 4481, 453–458.
- [70] Rajan Vaish, Q Vera Liao, and Victoria Bellotti. 2018. What’s in it for me? self-serving versus other-oriented framing in messages advocating use of prosocial peer-to-peer services. *International Journal of Human-Computer Studies*, 109, 1–12.
- [71] Marcus B Weaver-Hightower. 2017. Losing thomas & ella: a father’s story (a research comic). *Journal of Medical Humanities*, 38, 3, 215–230.
- [72] Leah Zhang-Kennedy, Robert Biddle, and Sonia Chiasson. 2017. Secure comics: an interactive comic series for improving cyber security and privacy. In *Proceedings of the 31st British Computer Society Human Computer Interaction Conference (HCI ’17)* Article 65. BCS Learning & Development Ltd., Sunderland, UK, 65:1–65:3.

## 8 APPENDIX

### 8.1 Posterior Predictive Check

In this section, we examine how well our model makes predictions on the observed data.

One of the advantages of a Bayesian model is that one can use the model to make predictions. Since the probabilistic model is generative, we can simply draw observations from the model. When we superimpose the posterior distribution traces onto the observational data in Figure 10 we can observe three things. First, notice that the mean contribution for the ‘text’ condition is less than the ‘comic’ or the ‘comic+social norm’ conditions, which seems reasonable given the increased number of \$0.0 contributions for the ‘text’ condition in comparison to the other two cases. Second, the use of the heavy-tailed  $t$ -distribution captures well the tails of the distribution at \$0 and at \$5.0. Third, the scale of the distribution for the text condition is smaller than the scales for the two comic conditions, justifying our use of different scale parameters  $\sigma_j$  for the three experimental conditions.

Having discussed posterior-prediction checks, next, we discuss alternatives to the model.

### 8.2 Alternative Models

As a first instance, consider a similar model, except that the scale parameter of the likelihood function is *not nested* like our current model. Instead, we consider the equal variance case, where all the variances are equal (similar to ANOVA), and that the variance is drawn from a uniform distribution. In other words,  $\sigma_j = \sigma \sim U(L, H)$ , where  $L > 0$  and where  $H$  is a large constant. The



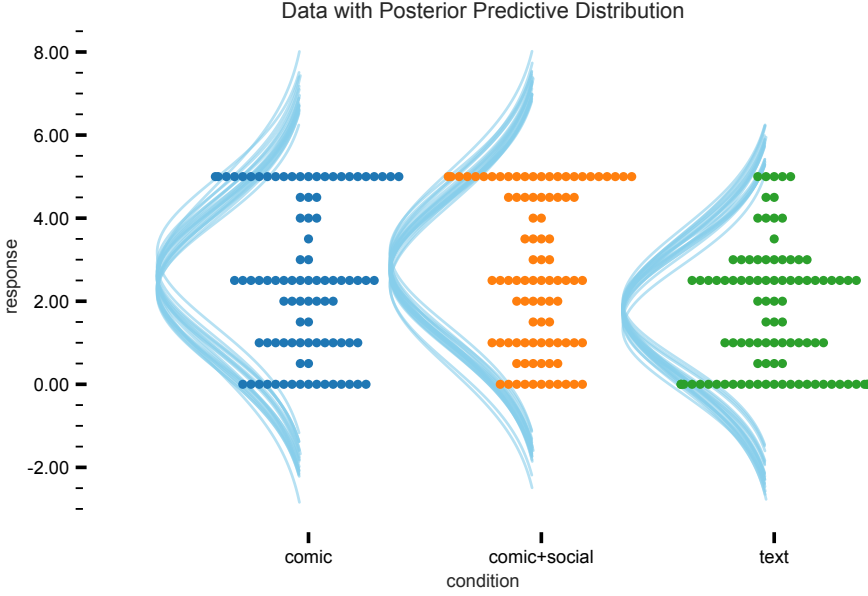


Fig. 10. Posterior prediction check: traces from the posterior distributions for each condition are super-imposed onto the observational data. The  $x$ -axis shows the three conditions, while the  $y$ -axis shows the contribution (each dot corresponds to a response by a single subject). The  $t$ -distribution captures the heavy tails, and we also see that the scale of the  $t$ -distribution in the case of the text condition is smaller than the other two cases.

main effect of the equal variance assumption is that there is no information sharing among the groups as would be the case when each  $\sigma_j$  is drawn from the same distribution, whose parameters have hyper-priors; the latter is our current model.

The main effect of constraining our simplified model is that we are slightly poorer in predicting the observed data since all the variances are guaranteed by the model to be equal, whereas we can see from Figure 6 that the mean scale (or equivalently variance) in the text condition is lower. We are skipping the traceplot and the contrast plots in this case, as they are similar, to Figure 6 and Figure 7, except that the effect size for the combined case is slightly lower due to the equal variance assumption.

Instead, we compare the two models using WAIC (Widely Applicable Information Criterion), a principled way to compare models when they have identical likelihood functions [25]. WAIC uses the predictive loss to compare two models with different parameters. First, WAIC computes the average log likelihood of each training data point (over the posterior distribution) less the variance of the log likelihood for the same data point; and then it computes the sum over all data points. That is, WAIC for a model:

$$\text{WAIC} = -2 \left( \sum_i^N \log \Pr(y_i) - V(y_i) \right)$$

where,  $N$  is the total number of training points,  $y_i$  is the observation,  $\Pr(y_i)$  is the averaged data likelihood over the posterior, and where  $V(y_i)$  is the variance of the data likelihood over the

posterior. When we compare the two models, one with unequal variances, and one with equal variances. We show our results in Table 1:

Table 1. WAIC comparison between the model with equal variances against the case when the variances are not constrained to be equal (i.e. we use a hierarchical model). Both cases assume the Student- $t$  likelihood function, the function used in this paper. The columns show respectively, WAIC, pWAIC (the effective number of parameters; also:  $pWAIC = \sum_i V(y_i)$ ), dWAIC (the difference between the WAIC scores of the other models with the best model), weight (the relative probability that the model explains the data) SE, the standard error of the WAIC estimate, dSE is the standard error of the difference of the current model against the top model. The table shows that the hierarchical model with unequal variances better explains the observations.

	Model	WAIC	pWAIC	dWAIC	weight	SE	dSE
Unconstrained variances, hierarchical		1102.48	3.95	0.00	1.00	14.61	0.00
Constrained, equal variances		1104.17	3.42	1.69	0.00	14.25	1.54

The results in Table 1 say that model with the unconstrained variances is better at explaining the data than the model with constrained variances; the relative probability that the hierarchical model with unconstrained variances better explains the observation is 1.0 (refer to the weight column in Table 1)

How useful was the choice of the Student- $t$  drawing distribution Equation (1), instead of assuming that the outcomes are drawn from a Normal distribution? Our analysis of the posterior distribution of the degrees of freedom parameter  $\nu$  shows that there is only a small probability ( $P(\nu \leq 30) \approx 0.07$ ) that  $\nu \leq 30$ . Thus, we may use the Normal likelihood function, without significantly affecting the conclusions. Indeed, in our experiments, when we do model the observations with a Normal likelihood, assume equal variances, we find no significant differences in the contrasts or the effect sizes, with the Student- $t$  model with equal variances (results omitted due to space constraints).

Since the charitable donations are bounded to lie between \$0.0 and \$5.0, might we benefit from using bounded likelihood functions like the Beta distribution  $Y \sim \text{Beta}(\alpha_j, \beta_j)$  to represent the charitable donations  $Y_j$  under the different conditions  $j$ ? While a  $\text{Beta}(\alpha, \beta)$  distribution lies between  $[0, 1]$ , we can scale down the contributions to lie in  $[0, 1]$  to use with the  $\text{Beta}(\alpha, \beta)$  distribution. But notice from Figure 4 that in each experimental condition, there is a central lobe, and heavy tails at each extreme, notably at \$0.0 and at \$5.0.

Received April 2019; revised Jun 2019; accepted Aug 2019