CENTER FOR
ECONOMIC
BEHAVIOR &
INEQUALITY

CEBI WORKING PAPER SERIES

Working Paper 27/20

THE VOLUNTEER'S DILEMMA EXPLAINS THE BYSTANDER EFFECT

Pol Campos-Mercade

ISSN 2596-44TX

CEBI
Department of Economics
University of Copenhagen
www.cebi.ku.dk

The Volunteer's Dilemma explains the Bystander Effect*

Pol Campos-Mercade†

November 27, 2020

Abstract

The *bystander effect* is the phenomenon that people are less likely to help others when they are in a group than when they are alone. The theoretical literature typically explains the bystander effect with the volunteer's dilemma: if providing help is equivalent to creating a public good, then bystanders could be less likely to help in groups because they free ride on the other bystanders. This paper uses a dynamic game to experimentally test such strategic interactions as an explanation for the bystander effect. In line with the predictions of the volunteer's dilemma, I find that bystanders help immediately when they are alone but help later and are less likely to help if they are part of a larger group. In contrast to the model's predictions, subjects in need of help are helped earlier and are more likely to be helped in larger groups. This finding can be accounted for in an extended model that includes both altruistic and selfish bystanders. The paper concludes that the volunteer's dilemma is a sensible way to model situations in which someone is in need of help, but it highlights the need to take heterogeneous social preferences into account.

Keywords: Volunteer's dilemma, Bystander effect, Helping behavior, Group size, Altruism.

JEL Codes: C92, D64, D90.

*I would like to acknowledge the Laboratory for Experimental Economics (LEE) and the Center for Economic Behavior and Inequality (CEBI) at the University of Copenhagen for financial support. I thank Jordi Brandts, David Cooper, Adrian Hillenbrand, Nagore Iriberri, Peter Jochumzen, Jesús Marín, Erik Mohlin, Daniele Nosenzo, Marina Núñez, Erik Wengström, the associate editor, and two anonymous referees for their helpful comments and suggestions. Special thanks to Marco Piovesan and Jesper Sam Sørensen for their guidance and engagement in the beginning of the project. I also thank participants at the 9th CNEE Workshop in Copenhagen, the IMEBESS Conference in Rome, the TRIBE meeting in Copenhagen, the M-BEES Conference in Maastricht, the ESA European Meeting in Bergen, the ESA North American Meeting in Tucson, the Behavioral and Organizational Design Workshop in Copenhagen, and seminar participants at Lund University and at the University of Barcelona. Earlier versions of this paper had the title "Helping Behavior and Group Size."

†Department of Economics, University of Copenhagen. *e-mail*: pcm@econ.ku.dk.

1 Introduction

People help one another even when it is not in their material interest. Passersby pay money to call an ambulance if someone requires medical attention. Internet users spend time answering strangers' questions. Car drivers stop to assist passengers with a damaged car. Social psychologists have studied such helping behavior for decades. One of the most robust conclusions from this stream of research is that a person's decision about whether to help depends crucially on the number of other people who also have the ability to help. More concretely, individuals who are part of a larger group are less willing to help (for a recent meta-analysis, see Fischer et al. 2011). Darley and Latané (1968) called this phenomenon the *bystander effect*.

An extensive literature has examined the reasons behind the bystander effect. The reviews by Latané and Nida (1981), Fischer et al. (2011), and Ross and Nisbett (2011) discuss the three main mechanisms that social psychologists generally use to explain this phenomenon. The first one is *evaluation apprehension*, where bystanders fear being negatively judged by others if they misinterpret the situation and there is no need for help. The second mechanism is *pluralistic ignorance*, which originates from relying on others' reactions to recognize the need for help in an ambiguous situation. That is, if the bystander sees other bystanders not helping, the bystander infers that there is no need for help. The third mechanism is *diffusion of responsibility*, which refers to the tendency to refrain from helping when others are present due to a reduced feeling of responsibility.

In contrast to these explanations, research in economics and sociology has mainly explained the bystander effect using the volunteer's dilemma model (Diekmann 1985). In the volunteer's dilemma, each player in a group simultaneously decides whether to volunteer to produce a public good. Those who volunteer pay a cost, and the public good is produced if at least one player volunteers. Since players free-ride on each other, in the only symmetric equilibrium the probability that each player volunteers decreases with group size. This literature notes that if bystanders help because they are altruistic, then providing help is similar to creating a public good. In this case, players' strategic interactions in the volunteer's dilemma—in which players are less likely to volunteer in larger groups—could explain the bystander effect observed in the field—in which bystanders are less likely to help in larger groups. While the strategic interactions predicted by the volunteer's dilemma are typically used to describe why people in larger groups are less likely to provide help (Archetti 2011; Bergstrom 2012, 2017; Bliss and Nalebuff 1984; Campos-Mercade 2020; Diekmann 1985, 1986, 1993; Franzen 2013; Fromell et al. 2019; Guha 2020; Harrington 2001; Healy and Pate 2018; Hillenbrand and Winter 2018; Hillenbrand et al. 2020; Kopányi-Peuker 2019; Tutic 2014; Przepiorka and Diekmann 2018; Weesie 1993, 1994), the connection between the volunteer's dilemma and the bystander effect lacks empirical evidence.[1]

In this paper, I test the strategic interactions predicted by the volunteer's dilemma as an explanation for the bystander effect. To do so, I first extend the volunteer's dilemma to a dynamic game. The dynamic setting is ideal because the model produces a set of testable predictions different from the ones implied by the mechanisms discussed in the psychology literature. More concretely, the model predicts that bystanders in larger groups will help later and be less likely to help. I then test these predictions in a lab experiment. Since bystanders in the experiment know exactly the extent to which the victim needs help, the experiment shuts down the evaluation apprehension and pluralistic ignorance mechanisms, which rely on whether bystanders interpret that the victim is in need of help (Ross and Nisbett 2011). The experimental results still show the bystander effect: subjects help later and are less likely to help when they are part of a larger group. While diffusion of responsibility, defined as "a lower psychological cost of not helping in larger groups" (Fischer et al. 2006), could explain why bystanders are less likely to help in larger groups, it cannot explain why bystanders help later.[2] The data thus support the free-riding on others implied by the volunteer's dilemma

[1] Note that the connection between the bystander effect and the volunteer's dilemma relies on two strong assumptions. First, if helping a victim is understood as producing a public good, then it must be that bystanders help because they care about the welfare of the victim—i.e., they are altruistic. In reality, however, bystanders' motive for helping could stem from other types of social preferences, such as warm-glow preferences (Andreoni 1990) or a desire to follow social norms (see, e.g., Krupka and Weber 2013). Second, bystanders are assumed to understand that other bystanders are also altruistic and to engage in strategic interactions with them about who will help. This requires bystanders to understand the situation as a game and to (correctly) infer the other bystanders' motives for helping. Due to the multiple confounds of the field, these assumptions are hard to test.

[2] Some papers in sociology and economics have used the term diffusion of responsibility to refer to the strategic interactions of the volunteer's dilemma. In this paper, I distinguish between both mechanisms: I use the definition of *diffusion of responsibility* used in the psychology literature and I use the term *strategic interaction* to refer to the free-riding produced by the volunteer's dilemma.

as an explanation for the bystander effect.

This result provides evidence that both the theoretical and the experimental results of the volunteer's dilemma are generalizable to situations in which someone needs help. The volunteer's dilemma can hence be used by policy-makers and organizations interested in understanding how to promote helping behavior. Additionally, it adds an additional mechanism, strategic interaction, to the channels considered by social psychologists to describe the bystander effect.

The paper begins by formulating the dynamic volunteer's dilemma, a game in which bystanders decide when to help a *victim* who loses utility over time. If bystanders are selfish, the model predicts that bystanders will never help. However, if bystanders are altruistic toward the victim, helping is equivalent to producing a public good. In this case, the model predicts that bystanders in a larger group help later and are less likely to help. Furthermore, the victim is also helped later and is less likely to be helped.

I find support for the predictions concerning bystanders' behavior using a lab experiment in which subjects in a group (bystanders) decide when to help another subject (the victim) who loses money over time. The first bystander(s) who helps pays a fixed cost and stops the victim from losing any further money. In line with the model, I find that when only one bystander can help the victim, most subjects help immediately. When two bystanders can help, most subjects shift their decision toward helping after several seconds, probably hoping for someone else to help before they do. When four bystanders can help, subjects wait even longer before helping.

In conflict with the predictions of the model, however, I find that *victims* in larger groups are helped earlier and are more likely to be helped.[3] I discuss different explanations for this discrepancy and argue that it is due to heterogeneity in subjects' willingness to help. By exploiting the within-subject dimension of the experiment, I find that a proportion of the subjects (approximately 30%) never help. I call these subjects *selfish bystanders*. I argue that the reason why victims in larger groups are better off is that the probability that all bystanders in the group are selfish is lower in a large group than in a small one.

[3]Since there are more bystanders who can help in larger groups, note that the fact that bystanders are less likely to help in larger groups does *not* imply that victims are less likely to be helped in larger groups.

The main contribution of this paper is to experimentally test whether the volunteer's dilemma explains the bystander effect. Previous research has merely noted that a model where helping is perceived as creating a public good could explain the bystander effect. However, social psychologists have found other alternative explanations for bystanders' behavior. I discuss them and argue that they cannot explain the bystander effect found in the experiment, concluding that strategic interaction is the main mechanism that drives the experimental results. Therefore, this paper gives external validity to the models that link the volunteer's dilemma with the bystander effect (e.g., Bergstrom 2012, 2017; Bliss and Nalebuff 1984; Diekmann 1985; Guha 2020; Hillenbrand and Winter 2018; Weesie 1993, 1994). Nevertheless, it calls for caution when applying the volunteer's dilemma to predict the probability that the victim is helped. These results strongly rely on the assumption that the bystanders are homogeneous. As in the lab, this assumption is unlikely to hold in reality.

Another contribution of this paper is to extend the volunteer's dilemma to a dynamic game and to study the effects of group size on volunteering over time. To the best of my knowledge, Bliss and Nalebuff (1984), Weesie (1993, 1994), and Otsubo and Rapoport (2008) are the only papers that have studied the volunteer's dilemma as a dynamic game. Bliss and Nalebuff (1984) and Weesie (1993, 1994) model a situation with continuous time, an infinite horizon, heterogeneous agents, and incomplete information. In the present paper, I show that a much simpler model can account for most of their results. The main advantage of this model is that it yields predictions that are more suitable for testing in the lab. Furthermore, its simplicity makes it easier to incorporate extensions.[4]

The paper also contributes to a literature that tests the volunteer's dilemma in the lab. Diekmann (1986) and Franzen (1995) show that subjects are less likely to volunteer in larger groups, although they overall volunteer more than predicted by the mixed strategy equilibrium. Goeree et al. (2017) corroborate these results and show that such behavior, which leads large groups to be more likely to produce the public good than smaller groups, is in line with the predictions of the quantal-response equilibrium (McKelvey and Palfrey

[4]As in the present paper, Otsubo and Rapoport (2008) model a game in discrete time and with a finite horizon. However, they study the effects of varying the costs to volunteer rather than group size. I extend their model by allowing agents to be unwilling to help in some periods, deriving expressions for the cumulative probabilities of helping, and studying the effects of group size on the bystanders' behavior and the victims' outcomes. In addition, Appendix B further extends the model to allow for two types of bystanders.

1995). More recently, Kopányi-Peuker (2019) and Hillenbrand et al. (2020) have tested the volunteer's dilemma with groups of up to 300 subjects. In conflict with the model, both papers find that subjects' behavior is insensitive to group size for such very large groups. Finally, in dynamic volunteer's dilemmas, Otsubo and Rapoport (2008) find that subjects volunteer earlier than predicted by theory, and Babcock et al. (2017) find that women are more likely to volunteer in mixed-gender groups.[5] In contrast to the present paper, this literature induces a value for the public good, such that all subjects receive a monetary reward if a subject volunteers. However, the applications and examples of the volunteer's dilemma very often involve implicit valuations which depend on each player's preferences. This paper presents a first test of the volunteer's dilemma's predictions in a setting in which the value of the public good is not an induced monetary reward.

Most previous experiments testing the bystander effect fall into the category of *natural field experiments* (as defined by Harrison and List 2004). While these kinds of experiments may have high external validity in their specific environment, the mechanisms are usually difficult to identify. This paper presents a novel experiment to test and study the bystander effect in a more controlled environment. To the best of my knowledge, only three recent papers have studied whether the bystander effect also exists in the lab. Panchanathan et al. (2013) find that dictators give less if there are other dictators who can also give. Fromell et al. (2019) show that dictators are more likely to give nothing when it implies being less likely to be chosen as the dictator. In a dynamic experiment in which recipients do not lose money over time and only one dictator can be the helper, Bergstrom et al. (2019) find that dictators in larger groups help later and are less likely to help. These papers carefully analyze the helping motives of bystanders who are in a group with others. In contrast, the present paper contributes to this literature by studying the reasons why the bystander effect emerges—this is, the reasons why people are less likely to help in larger groups. More concretely, the paper replicates the bystander effect and shows that it is generated by the strategic interactions predicted by the volunteer's dilemma.

The paper proceeds as follows. Section 2 introduces the dynamic volunteer's dilemma

[5] See also Diekmann (1993) for evidence on the volunteer's dilemma with asymmetric payoffs and Healy and Pate (2018) and Hillenbrand and Winter (2018) for evidence on the volunteer's dilemma with incomplete information.

game and presents six experimental hypotheses. Section 3 outlines the experimental design used to test the hypotheses. Section 4 presents the experimental results. Section 5 discusses the mechanisms behind the results. Section 6 concludes.

2 The dynamic volunteer's dilemma game

This section formulates a model in which a group of bystanders can help a victim over a finite horizon in discrete time. While a static volunteer's dilemma would yield simpler predictions, in the experiment it would not be possible to disentangle whether bystanders in larger groups do not help because they are no longer willing to pay the cost (diffusion of responsibility) or because they are willing but hope that someone else does it first (strategic interaction). Exploiting the dynamic dimension helps teasing out these two mechanisms because strategic interaction does not only predict that bystanders are less likely to help but also that they help later.[6] Moreover, the dynamic game is interesting in its own right. Many (maybe most) situations in which someone is in need of help have some time dimension. Hence, it is not only interesting to study whether people help, but also when they help.

The first part of this section describes the setup of a game. The second part assumes that the bystanders are homogeneous and altruistic toward the victim. This transforms the game into a dynamic volunteer's dilemma with a finite horizon, discrete time, and homogeneous agents.[7] I then derive the (only) symmetric equilibrium to form testable hypotheses.

2.1 Setup of the game

Every bystander in a group of $n \geq 1$ bystanders can assist the victim. Time is discrete and finite. The victim loses utility over time across $T + 1$ periods, $0, 1, ..., T$. In each period,

[6] In fact, the numerical simulations in Appendix C show that, in the specific setting of the experiment, strategic interaction mainly predicts treatment differences in *when* bystanders help, rather than *whether* they help.

[7] A model with different types of bystanders, incomplete information, and a finite horizon becomes unnecessarily complicated and often lacks analytical solutions. Appendix B extends the model to include two types of bystanders and complete information. It shows that the results for the bystanders' behavior hold but that those concerning the victim's outcome do not. Section 5 discusses in detail the implications of such differences for both models. The results obtained in this section and in Appendix B turn out to be very similar to those obtained by Bliss and Nalebuff (1984) and Weesie (1993), who rely on different assumptions.

every bystander decides simultaneously and independently to help or not to help. The game starts at $t = 0$ and terminates either when at least one bystander helps at $t < T$ or at T if no bystander helps before then. Hence, a (mixed) strategy for a bystander specifies, for each period, the probability that the bystander helps in that period, conditional on no one having helped before. The first bystander who decides to help is called the *helper* and pays a fixed cost. If several bystanders help in the same period, then these bystanders become the helpers and pay the cost.

2.2 Payoffs

If bystanders were purely selfish, they would never pay any positive cost to help someone who cannot reciprocate. In this game, this implies that bystanders would not help in any period and would always wait for the game to end at period T (i.e., the only subgame perfect equilibrium for each bystander is to not help in any period). However, the volunteer's dilemma is often used to model situations in which someone is in need of help. Following this literature, I assume that the utility that bystanders obtain if the victim is immediately helped is higher than the cost of helping. Furthermore, because the victim loses payoff over time, I assume that bystanders prefer that the victim be helped earlier rather than later. However, because helping carries a cost, each bystander prefers someone else to help. This setup is equivalent to a dynamic volunteer's dilemma game in which producing the public good is helping the victim.

Denote by H_t and NH_t the respective implicit utilities of the helper and the non-helper if the game ends at period t. Assume $H_t > H_{t+1}$ and $NH_t > NH_{t+1}$ for every $t \in \{0, 1, \ldots, T-1\}$. This captures the fact that bystanders prefer that victims are helped earlier than later. To reflect the cost of helping, assume $NH_t > H_t$ for every $t \in \{0, 1, \ldots, T-1\}$. If the game reaches period T (that is, no one ever helps), let the bystanders' payoff be NH_T. Assume $H_0 > NH_T$, implying that bystanders prefer to help at $t = 0$ rather than that the victim is never helped.[8]

[8]Note that the model does not rule out that bystanders also have warm-glow preferences; it assumes instead that the warm-glow utility of helping is lower than the cost of helping.

2.3 Equilibrium

Since players are identical (and since subjects cannot communicate in the experiment below), the analysis is restricted to symmetric subgame perfect equilibria in behavioral strategies (see, e.g., Fudenberg and Tirole 1991). In this equilibrium, players help in each period with a given probability. However, they stop helping after a given period. The reason is that from that period onwards, the utility that they gain if the victim is helped becomes lower than the cost of helping.

Proposition 1. *If $H_{T-1} < NH_T$, let $\tau \in \{1, \ldots, T-2\}$ be the last period in which $H_t \geq NH_T$. Otherwise, let $\tau = T-1$. In the unique symmetric subgame perfect equilibrium, bystanders help with probability σ_t if $t < \tau$, σ_τ if $t = \tau$, and 0 if $t > \tau$, where*

$$\sigma_t = 1 - \left(\frac{NH_t - H_t}{NH_t - H_{t+1}} \right)^{\frac{1}{n-1}} \tag{1}$$

and

$$\sigma_\tau = 1 - \left(\frac{NH_\tau - H_\tau}{NH_\tau - NH_T} \right)^{\frac{1}{n-1}}. \tag{2}$$

The proof is in Appendix A.

Note the similarity between the above equilibrium and the equilibrium in the static volunteer's dilemma, where the probability to volunteer is one minus the $(n-1)^{th}$ root of the cost-benefit ratio. For σ_t, the expression in the numerator, $NH_t - H_t$, captures precisely the cost of helping. Moreover, the expression in the denominator, $NH_t - H_{t+1}$, captures the benefit that someone helps in period t. This is because, since bystanders play mixed strategies in each period, the expected utility of each bystander if the game moves on to period $t+1$ is H_{t+1}.[9] Hence, this equilibrium can be thought of as bystanders engaging in a static volunteer's dilemma in each period.

2.4 Hypotheses

In this subsection, I use the mixed strategy equilibrium associated with the above equilibrium in behavioral strategies to make hypotheses about bystanders' behavior. Section 4 contrasts

[9]For σ_τ, since τ is the last period in which bystanders are willing to help, the expected utility is NH_T if no one helps.

these predictions with the experimental data. Appendix A contains the calculations and analyses.

Consider $n = 1$. Since $H_0 > NH_T$ and $H_t > H_{t+1}$, the model predicts that all bystanders help in period 0.

Hypothesis 1. *When only one bystander can help the victim, the bystander will help immediately.*

When $n \geq 2$, bystanders play as in the equilibrium above. That is, they help at each period $t \in \{0, 1, \ldots, \tau\}$ with probability σ_t and do not help for $t > \tau$. Therefore, bystanders no longer only help immediately.

Hypothesis 2. *When more than one bystander can help, bystanders may help after several periods.*

Will bystanders in larger groups help earlier or later? To answer this question, I compute the cumulative probability that a bystander will help in period t, conditional on no bystander having helped before. This probability decreases with n, which implies that bystanders in larger groups help later on average.

Hypothesis 3. *Each bystander helps later on average in larger groups.*

In a similar way, the classical result of the static volunteer's dilemma, in which bystanders in larger groups are more likely not to help, holds.

Hypothesis 4. *Each bystander's probability of not helping is higher in larger groups.*

One can also analyze the problem from the victim's perspective. Note that the fact that bystanders in larger groups help later is not sufficient to conclude that victims are helped later in larger groups. When group size increases, two opposite effects occur. First, the victim is less likely to be helped earlier because bystanders decide to help later. Second, the victim is more likely to be helped earlier because there are additional bystanders who can help early. In this model, it turns out that the first effect dominates the second one.[10]

Hypothesis 5. *Victims in larger groups are on average helped later. Furthermore, victims in larger groups are less likely to be helped.*

[10] Appendix B shows that this prediction does not hold when there can be two types of bystanders.

3 Experimental design

This section describes the design of the experiment used to test the theoretical hypotheses. The experimental data were collected at the Laboratory for Experimental Economics (LEE) at the University of Copenhagen. Eighty subjects were recruited using ORSEE (Greiner 2004) and participated in an experiment programmed in z-Tree (Fischbacher 2007). There were four sessions that lasted approximately 90 minutes and consisted of twenty subjects each. Subjects participated in the experiment anonymously using different computers and without communicating with one another. They earned an average of 131 DKK ($\sim$ €18), of which 50 DKK corresponded to a participation fee.

In the beginning of the session, the computer randomly assigned each subject to the role of either bystander or victim (labeled active participant and passive participant), and this role was kept fixed throughout the experiment. In each round, the computer matched a victim to a group of one, two, or four bystanders. Each bystander started with an endowment of 24 points. The victim's payoff started at 20 points and decreased one point every second until a bystander helped her or until 20 seconds had passed. The first bystander who helped became *the helper* and paid a fixed cost of 4 points.

The experiment consisted of two parts of seven rounds each. In each part, subjects participated (in random order) in the one-bystander treatment once, in the two-bystander treatment twice, and in the four-bystander treatment four times.[11] Subjects were informed that the bystanders and the victim in their group would change after each round.[12]

In the first seven rounds of the experiment, which are the ones mainly used in the analysis below, each bystander selected the number of seconds that he was willing to wait before helping, if no one helped before (as illustrated in Figure 1). The bystander(s) who selected the lowest number of seconds became the helper and paid a fixed cost of 4 points,

[11] This matching allowed all subjects to play in all rounds and to participate at least once in each treatment in each part.

[12] Most of the previous literature on the bystander effect uses between-subject experimental designs. There are two main advantages of using a within-subject design in this paper. First, it greatly increases the statistical power of the experiment, since the intra-subject correlation of waiting time to help is relatively high (0.65). Second, and most importantly, this within-subject design can be used to classify subjects, which turns out to be crucial for understanding whether the model correctly predicts their behavior. In fact, Section 5 rules out several alternative explanations of the results by using this feature of the experiment.

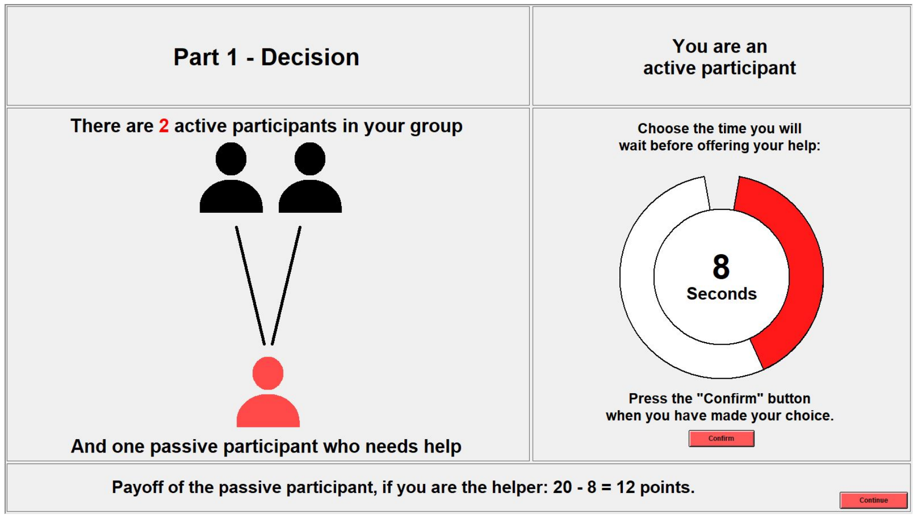

Figure 1. Screen to select the number of seconds to wait before helping when two bystanders can help.

while the rest did not pay any cost. For example, if in the two-bystander treatment the first bystander chose to wait 2 seconds and the second bystander chose to wait 5 seconds, the victim would be helped after 2 seconds (and earn 20 - 2 = 18 points), and the first bystander would pay the cost of 4 points. The victim was thus helped after the number of seconds that the helper chose. If a bystander never wanted to help, then he had to select to wait 20 seconds. If all bystanders in a group selected never to help, then the victim lost her 20 points, and no one paid the cost of helping. Only the victim and the subject who was the helper saw whether and when the victim was helped.[13]

In the last seven rounds, of which the subjects had no information beforehand, subjects played in real time rather than using the strategy method. Hence, instead of selecting the number of seconds to wait before helping, bystanders saw a clock ticking in real time and could choose to stop it at any given moment. The clock stopped for all bystanders once one bystander stopped it. The first bystander who stopped the clock paid the cost of helping, and the victim was helped at that moment.

The real-time method likely increases the realism of the helping situation. However, its

[13]Since the experiment was relatively long, the feedback was given to prevent subjects' answers from deteriorating due to boredom (Maniaci and Rogge 2014). Although this feedback could potentially lead to asymmetric information (the bystander who helped knows that no one helped before him, while the rest only know that another bystander helped at some point before they decided to help), the regression estimates remain constant when controlling for whether subjects helped in the previous one or two rounds.

drawback is that one can only observe when a victim is helped, rather than when each of the bystanders would have chosen to help. This makes the method less suitable to study the predictions of the model. I hence use the first seven rounds to test the models' hypotheses, and use these last seven rounds as a robustness check to confirm that the choices made using the strategy method are not systematically different to those made in real time (e.g., Brandts and Charness 2011).

The final payment consisted of the sum of each subject's earnings for each of the fourteen rounds (at the rate 4 points = 1 DKK).[14]

4 Results

Table 1 presents the descriptive statistics for the number of seconds that bystanders choose to wait before helping and the number of seconds until victims are helped. As expected, bystanders help later as the group size increases. However, contrary to the model's predictions, victims are helped earlier in larger groups than in smaller groups. In what follows, I first analyze bystanders' decision to help and then study when the victims are helped.

4.1 Bystander's perspective

Figure 2 illustrates the cumulative frequencies of the number of seconds that bystanders wait before helping. When only one bystander can help the victim, in contrast to the model in Section 2, a fraction (25%) of the subjects never help. However, as predicted by *Hypothesis 1*, most subjects (60.7%) help immediately. This suggests that those subjects behave altruistically: they do not only care about helping but also want the victim's payoff to be as high as possible.[15]

[14]The decision to pay all rounds, rather than one randomly drawn round, is to prevent subjects from artificially helping due to warm glow. Stahl and Haruvy (2006) suggest that paying fewer rounds could increase altruistic behavior since warm glow could be felt regardless of whether a round is implemented. In this experiment, such artificial warm glow could make subjects more likely to help immediately regardless of the treatment and thus bias the experimental results. While the drawback of paying all rounds is that subjects could theoretically hedge (Azrieli et al. 2018; Holt 1986), there is little evidence for such portfolio effects in experiments involving social preferences (Charness et al. 2016).

[15]Interestingly, 9% of the subjects help either in the first or in the second period, rather than helping immediately. This result could be because those subjects are altruistic but have some *desire to be ahead* (Fershtman et al. 2012).

Table 1. Average time at which bystanders help and victims are helped by treatment

	Bystander helps	Victim is helped
One Bystander	5.55	5.55
	(8.74)	(8.74)
Two Bystanders	9.45	4.21
	(8.67)	(6.53)
Four Bystanders	10.97	3.20
	(8.50)	(4.85)

Note: This table presents the averages of the number of seconds that bystanders select to wait before helping and the number of seconds until victims are helped by group size. Bystanders who do not help are weighted as waiting 20 seconds. These statistics pool all sessions and rounds across all subjects. Standard deviations are in parentheses.

Result 1. *When only one bystander can help the victim, that bystander either helps immediately or does not help.*

Note that as soon as the groups become larger than one bystander, the frequency of bystanders immediately helping is reduced from 60.7% to 17.9% for the two-bystander treatment and to 11.1% for the four-bystander treatments. As represented in Figure 2, this reduction is mainly explained by bystanders helping later rather than not helping at all. As predicted by *Hypothesis 2*, subjects spread their decision to help over time when groups are larger than one bystander.

Result 2. *When more than one bystander can help, most bystanders delay the time at which they help.*

Figure 2 shows that the cumulative probability that a bystander helps after a given number of seconds decreases with group size. To study the extent to which helping time varies across treatments, Table 2 uses a semiparametric Cox proportional hazard model with subject fixed effects, where the threshold is the one-bystander treatment. Column 4 shows that when controlling for each session and the round within each session, the hazard of selecting to help at any point in time is 61.9% lower in the two-bystander treatment than in the one-bystander treatment ($p = 0.001$) and 93% lower in the four-bystander than in the

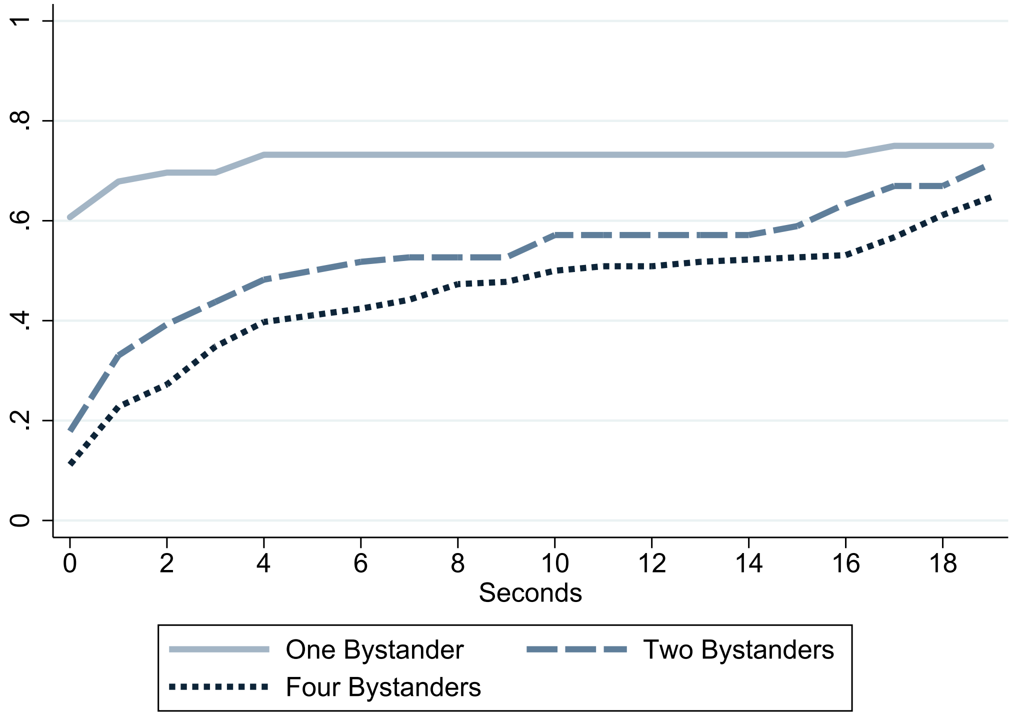

Figure 2. Cumulative frequency of bystanders selecting to help at each second by group size.

one-bystander treatment ($p < 0.001$). Furthermore, the coefficients for the two-bystander treatment and the four-bystander treatment are significantly different ($p = 0.013$).[16] The controls for sessions and rounds within each session are not significant. Hence, as predicted by *Hypothesis 3*, bystanders in larger groups help later.

Result 3. *The cumulative probability that a bystander helps at any given period decreases with group size. This implies that bystanders in larger groups on average help later.*

Figure 2 shows that the proportion of bystanders who decide to never help increases with group size, as predicted by *Hypothesis 4*. In particular, the proportion of bystanders who never help is 25% in the one-bystander treatment, 28.6% in the two-bystander treatment, and 35.3% in the four-bystander treatment. To test whether these differences are statistically significant, Column 4 in Table 3 uses an OLS regression with subject fixed effects in which the outcome variable is whether the subject chooses to help at some point—i.e., whether he

[16]The coefficients are very similar when using a random effects specification and when adding dummies that capture whether the subject helped in the previous round.

Table 2. Cox proportional hazard regression about the decision to help after a specific number of seconds

	(1)	(2)	(3)	(4)
Two Bystanders	-0.295*	-0.330*	-0.566***	-0.619***
	(0.164)	(0.172)	(0.176)	(0.181)
Four Bystanders	-0.433***	-0.475***	-0.878***	-0.930***
	(0.150)	(0.157)	(0.164)	(0.168)
Observations	392	392	392	392
Mean seconds	9.763	9.763	9.763	9.763
Round*Session FE		X		X
Subject FE			X	X

Note. This table reports the results of a semiparametric Cox proportional hazard model to test whether the helping times vary across treatments. The second and fourth columns include subject fixed effects (Subject FE). The third and fourth columns include dummy controls for session and round within each session (Round*Session FE). Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

chooses 19 seconds or earlier. Subjects in the one-bystander treatment are not significantly more likely to help at some point than in the two-bystander treatment ($p = 0.443$), but they are with respect to the four-bystander treatment ($p = 0.016$). Furthermore, subjects in the two-bystander treatment are more likely to help than in the four-bystander treatment ($p = 0.042$). Hence, while the differences are not always statistically significant, they seem to move in the direction predicted by the model.

Result 4. *Bystanders in larger groups choose never to help more often (the differences are not always statistically significant).*

Interestingly, some subjects in the two- and four-bystander treatments decide to help after 16 to 19 seconds if no one has helped yet. At that point, bystanders pay a cost of 4 points to help the victim earn 1 to 4 points. A model based on altruistic agents cannot successfully explain this behavior. One potential explanation is related to issues of self-image: these subjects do not want to see themselves as someone who would never help. They therefore choose to help at a very late point to preserve their self-image while hoping for someone else to help first.[17] Indeed, the great majority of them end up not helping, since someone in their

[17]A second explanation is that these subjects do not understand the game. However, this explanation

Table 3. OLS regression on the probability of offering help

	(1)	(2)	(3)	(4)
Two Bystanders	-0.036	-0.036	-0.036	-0.036
	(0.076)	(0.077)	(0.046)	(0.046)
Four Bystanders	-0.103	-0.103	-0.103**	-0.103**
	(0.070)	(0.070)	(0.042)	(0.042)
Observations	392	392	392	392
Mean probability	0.681	0.681	0.681	0.681
Round*Session FE		X		X
Subject FE			X	X

Note. This table reports the results of an OLS regression to test whether the probability of helping—i.e., a dummy where 1 corresponds to selecting 19 or fewer seconds—differs by treatment. The second and fourth columns include subject fixed effects (Subject FE). The third and fourth columns include dummy controls for session and round within each session (Round*Session FE). The results are equivalent in terms of significance when using probit and logit regressions. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

group chooses to help earlier.

4.2 Victim's perspective

Both the theoretical model and the experimental results indicate that bystanders in larger groups help later and are less likely to help. This section studies the victim's perspective and aims at answering how victims fare depending on the group size.

Figure 3 shows the cumulative probability that a victim is helped at any given period with respect to group size. While victims are more likely to be helped immediately in the one-bystander treatment, they are less likely to be helped before 3 seconds. In fact, victims in larger groups are on average helped earlier.

I use a Wilcoxon test for equality of survival functions to test for differences in when victims are helped across treatments. None of the differences turns out to be significant ($p =$

does not seem plausible. A closer examination of the data (not reported) suggests that subjects who help in the last periods do understand other factors of the game. All of them, for example, help either immediately or never when they are the only ones who can help (which is to be expected from most models of social preferences).

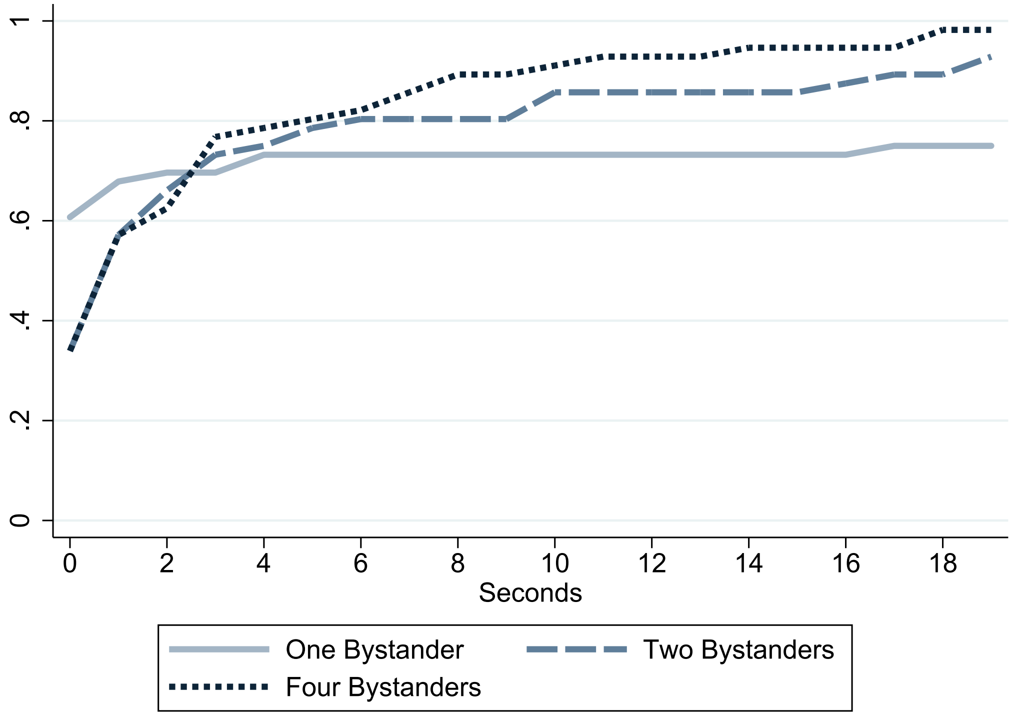

Figure 3. Cumulative frequency of the victim being helped at each second by group size.

0.233 when comparing the one- and two-bystander treatments, $p = 0.173$ when comparing the one- and four-bystander treatments and $p = 0.834$ when comparing the two- and four-bystander treatments). Cox proportional hazard regressions (not reported) also reveal that victims in larger groups are helped earlier on average, but the differences are not significant.

The proportion of victims who are helped is 75% in the one-bystander treatment, 92.9% in the two-bystander treatment, and 98.2% in the four-bystander treatment. A proportion test to compare frequencies shows that victims are significantly less likely to be helped in the one-bystander treatment than in the two-bystander treatment ($p = 0.010$) and four-bystander treatment ($p < 0.001$). Victims are also more likely to be helped in the four-bystander treatment than in the two-bystander treatment, but this difference is not significant ($p = 0.170$).

Result 5. *Contrary to the model's predictions, victims in larger groups are on average helped earlier (not significant). Furthermore, they are also more likely to be helped.*

All these results qualitatively and quantitatively replicate when analyzing subjects' behavior in the last seven rounds, in which the game is played in real time. In these rounds, victims are helped after 5.34 seconds in the one-bystander treatment, 5.18 seconds in the two-bystander treatment, and 2.46 seconds in the four-bystanders treatment. Similarly, the proportion of victims who are helped is 78.6% in the one-bystander treatment, 87.5% in the two-bystander treatment, and 92.2% in the four-bystander treatment. Hence, regardless of the elicitation method, and in stark contrast with *Hypothesis 5*, victims are better off in larger groups.

5 Discussion

This section discusses the two main results of the experiment. First, it discusses how bystanders' decision to help changes with group size. Second, it discusses some of the potential explanations for why victims are better off with larger groups.

5.1 Why do bystanders in larger groups help later?

The results for bystanders' behavior fit the hypotheses derived from the dynamic volunteer's dilemma, which assumes that bystanders perceive helping as a public good. Indeed, most bystanders help immediately when they are the only ones who can help. However, as soon as other bystanders can also help, bystanders both help later and are less likely to help. In what follows, I argue that this result is most likely due to the strategic interaction captured in the volunteer's dilemma, rather than the other explanations used in the psychology literature. Importantly, this does not imply that such other explanations are unimportant to generate the bystander effect in other settings. Rather, the results in this paper add an additional channel, strategic interaction, to those previously considered.

The experiment in this paper rules out the *evaluation apprehension* and *pluralistic ignorance* mechanisms considered in the psychology literature, which rely on whether bystanders interpret that the victim is in need of help (Ross and Nisbett (2011)). In the experiment, bystanders know exactly the extent to which the victim is in need of help. In particular, there is no reason to believe that bystanders consider the victim to be in less need in a larger group.

An alternative explanation for the results could be the *diffusion of responsibility* mechanism. In that case, bystanders in larger groups feel less responsible for the victim and therefore shift their decision from helping in small groups to not helping at all in large groups. This may be predicted by Fischer et al. (2006)'s definition of diffusion of responsibility, which is that bystanders feel a lower psychological cost of not helping when they are in larger groups.[18] In line with this mechanism, bystanders in the experiment are less likely to help when they are in larger groups. However, the main treatment difference is that bystanders in larger groups decide to help later, rather than not helping at all. This result seems to indicate that bystanders in larger groups do still take responsibility for helping the victim, but they just hope that someone else will help before they do. This is precisely the kind of strategic interaction that the volunteer's dilemma captures.

5.2 Why are victims more likely to be helped by larger groups?

The theoretical model predicts that victims are helped later and are less likely to be helped by larger groups. However, the experimental results conflict with this prediction: victims are helped earlier and are more likely to be helped by larger groups. This section exploits the within-subject dimension of the experiment to explore three potential reasons why the model fails to predict victims' outcomes: bystanders' warm-glow preferences, mistakes, and heterogeneity in preferences.

First, it may be that some subjects help due to warm-glow preferences (Andreoni 1990). For example, Bergstrom et al. (2019) find that between 15% and 36% of their subjects help because they want to be the helpers themselves (a similar behavior could arise if subjects had maximin preferences instead, see Guha 2020). If most subjects help due to warm-glow preferences, this could indeed explain that victims are better off with larger groups. For example, if 25% of the sample helped due to warm-glow preferences, then the probability of being helped by one of those subjects would be 25% in the one-bystander treatment, 44% in the two-bystander treatment, and 68% in the four-bystander treatment. Interestingly, this is not what the data show. Victims are still more likely to be immediately helped in the one-

[18] In fact, Cryder and Loewenstein (2012) find in the lab that the feeling of responsibility plays a role in dictators' decision to give to others (see also Dana et al. 2007).

bystander treatment (60.7%) than in the two- and four-bystander treatments (33.9%). Only 11.2% of the decisions in the four-bystander treatment are to help immediately, although a closer examination of the data shows that not a single bystander in the sample decides to always help immediately. All the subjects “bargain” about who will help the victim in at least one period. It is thus unlikely that the main explanation for why victims are better off with larger groups is the existence of warm-glow helpers who always help immediately.

Second, victims may be better off with large groups because bystanders make mistakes when deciding when to help. If subjects make mistakes (such as in quantal response equilibria, McKelvey and Palfrey 1995), then it is easier for at least one subject to help early by mistake in larger groups than in smaller groups. Allowing for mistakes in the model would preserve the results obtained at the bystander level, but victims could end up receiving more help with larger groups of bystanders. However, a closer look at the data suggests that mistakes were not especially common. Take, for example, the bystanders' decision in the one-bystander treatment. According to almost every social preferences model, these bystanders should help either immediately or never, so there are no a priori reasons why they would help at some other point in time. Indeed, 86% of the subjects help immediately or never (an additional 9% help during the first two periods). Assuming that these 14% of the choices are mistakes and that this proportion does not differ across treatments, such a low proportion of mistakes cannot account for the fact that victims are better off with larger groups.[19]

Third, victims might be better off with larger groups because some bystanders are never willing to help. For example, assume that a fraction of the bystanders are altruistic toward the victim (i.e., they would be willing to pay the cost to help her if no one else helped) and a fraction of the bystanders are not (i.e., they would never help). In this case, larger groups may be more beneficial for victims because they are more likely to contain at least one altruistic bystander.

This heterogeneity in preferences seems the most reasonable explanation of the results. In every treatment, approximately 30% of the bystanders never help (25%, 29%, and 35%

[19]Furthermore, subjects' decisions are relatively consistent across rounds. The average within-subject standard deviation of the choices made in the two- and the four-bystander treatments are 2.25 and 3.16 seconds, respectively. This suggests that the answers are not particularly noisy.

in the one-, two-, and four-bystander treatments). Further analysis of the data shows that these bystanders are consistent across treatments. For example, 87% of the decisions made by subjects who do not help in the one-bystander treatment are to not help in the other treatments either. In contrast, only 15.87% of the decisions made by subjects who help at some point in the one-bystander treatment are to not help in the other treatments. Now, label the bystanders who consistently do not help across treatments as *selfish bystanders* and assume that they account for 30% of the bystanders. Note that while the probability that the victim is assigned to a group with only selfish bystanders in the one-bystander treatment is 30%, this probability becomes 9% in the two-bystander treatment and 0.8% in the four-bystander treatment. Therefore, while it is quite likely that a victim is assigned to a group where no one is willing to help her in the one-bystander treatment, the chances that no one wants to help her in the larger groups are greatly reduced. The data show that, indeed, 25%, 8%, and 1% of the victims are not helped in the one-, two-, and four-bystander treatments, respectively. These percentages are very close to the discussed estimates of 30% selfish bystanders. According to this explanation, these are the victims who had the bad fortune of being assigned to a group with only selfish bystanders.

Appendix B extends the theoretical model by allowing a proportion of the bystanders to be selfish. The extended model makes the same predictions for bystanders' equilibrium behavior (*Hypotheses 2*, *3*, and *4*) but different predictions regarding the outcomes for the victims (*Hypothesis 5*). More precisely, the model predicts that victims will be more likely to be helped by larger groups when the fraction of selfish bystanders is sufficiently large. Appendix C presents a simulation exercise and shows that the extended model can also predict subjects' behavior in quantitative terms.

6 Conclusion

The bystander effect is the phenomenon that people are less likely to help when they are in a group than when they are alone (Darley and Latané 1968). In this paper, I find support for the hypothesis that the bystander effect emerges because bystanders strategically interact about who will help, as hypothesized by the volunteer's dilemma.

This result suggests that previous (and future) results on the volunteer's dilemma could be generalizable to situations in which someone needs help. One-way communication by the bystanders or taking turns helping might thus increase the probability that the victim is helped (Feldhaus and Stauf 2016 and Leo 2017); uncertainty about the number of bystanders or their cost of helping might increase helping rates (Hillenbrand and Winter 2018 and Healy and Pate 2018); and in situations in which both men and women are equally capable of providing help, women might be more likely to help than men (Babcock et al. 2017). These conjectures open new and exciting avenues for future research.

One should, however, be aware that a model with homogeneous bystanders may be too simplistic in some contexts. In this paper, although the model predicts bystanders' behavior relatively well, it does not correctly predict that victims are more likely to be helped and are helped earlier with larger groups. This is because there are bystanders who are just never willing to help. Because larger groups are less likely to contain only such selfish bystanders, victims are more likely to be helped. This insight leads to an interesting conjecture: victims should prefer smaller groups in situations in which almost all bystanders are willing to help them. However, victims should prefer larger groups in situations in which the proportion of selfish bystanders is sufficiently large.

References

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401):464–477.

Archetti, M. (2011). A strategy to increase cooperation in the volunteer's dilemma: reducing vigilance improves alarm calls. *Evolution: International Journal of Organic Evolution*, 65(3):885–892.

Azrieli, Y., Chambers, C. P., and Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4):1472–1503.

Babcock, L., Recalde, M. P., Vesterlund, L., and Weingart, L. (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review*, 107(3):714–47.

Bergstrom, T. (2017). The good samaritan and traffic on the road to jericho. *American Economic Journal: Microeconomics*, 9(2):33–53.

Bergstrom, T., Garratt, R., and Leo, G. (2019). Let me, or let george? motives of competing altruists. *Games and Economic Behavior*, 118:269–283.

Bergstrom, T. C. (2012). Single-payer volunteers' dilemma. *UCSB Working Paper*.

Bliss, C. and Nalebuff, B. (1984). Dragon-slaying and ballroom dancing: The private supply of a public good. *Journal of Public Economics*, 25(1):1–12.

Brandts, J. and Charness, G. (2011). The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics*, 14(3):375–398.

Campos-Mercade, P. (2020). When are groups less moral than individuals? *Working paper*.

Charness, G., Gneezy, U., and Halladay, B. (2016). Experimental methods: Pay one or pay all. *Journal of Economic Behavior & Organization*, 131:141–150.

Cryder, C. E. and Loewenstein, G. (2012). Responsibility: The tie that binds. *Journal of Experimental Social Psychology*, 48(1):441–445.

Dana, J., Weber, R. A., and Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1):67–80.

Darley, J. M. and Latané, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4):377–383.

Diekmann, A. (1985). Volunteer's dilemma. *Journal of Conflict Resolution*, 29(4):605–610.

Diekmann, A. (1986). Volunteer's dilemma. a social trap without a dominant strategy and some empirical results. *Paradoxical Effects of Social Behavior*, pages 187–197.

Diekmann, A. (1993). Cooperation in an asymmetric volunteer's dilemma game theory and experimental evidence. *International Journal of Game Theory*, 22(1):75–85.

Feldhaus, C. and Stauf, J. (2016). More than words: the effects of cheap talk in a volunteer's dilemma. *Experimental Economics*, 19(2):1–18.

Fershtman, C., Gneezy, U., and List, J. A. (2012). Equity aversion: Social norms and the desire to be ahead. *American Economic Journal: Microeconomics*, 4(4):131–44.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fischer, P., Greitemeyer, T., Pollozek, F., and Frey, D. (2006). The unresponsive bystander: Are bystanders more responsive in dangerous emergencies? *European Journal of Social Psychology*, 36(2):267–278.

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., Moritz, H., Wicher, M., and Kainbacher, M. (2011). The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, 137(4):517–537.

Franzen, A. (1995). Group size and one-shot collective action. *Rationality and Society*, 7(2):183–200.

Franzen, A. (2013). The volunteer's dilemma: Theoretical models and empirical evidence. In *Resolving social dilemmas: Dynamic, structural, and intergroup aspects*. Psychology Press.

Fromell, H., Nosenzo, D., Owens, T., and Tufano, F. (2019). Are victims truly worse off in the presence of bystanders? revisiting the bystander effect. *Revue économique*, 70(6):927–943.

Fudenberg, D. and Tirole, J. (1991). *Game theory*. The MIT Press.

Goeree, J. K., Holt, C. A., and Moore, A. K. (2017). An experimental examination of the volunteer's dilemma. *Games and Economic Behavior*, 102:303–315.

Greiner, B. (2004). An online recruitment system for economic experiments. In *Forschung und Wissenschaftliches Rechnen*, pages 79–93. Kurt Kremer and Volker Macho.

Guha, B. (2020). Revisiting the volunteer's dilemma: group size and public good provision in the presence of some ambiguity aversion. *Economics Bulletin*, 40(2):1308–1318.

Harrington, J. E. (2001). A simple game-theoretic explanation for the relationship between group size and helping. *Journal of Mathematical Psychology*, 45(2):389–392.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic literature*, 42(4):1009–1055.

Healy, A. J. and Pate, J. G. (2018). Cost asymmetry and incomplete information in a volunteer's dilemma experiment. *Social Choice and Welfare*. https://doi.org/10.1007/s00355-018-1124-6.

Hillenbrand, A., Werner, T., and Winter, F. (2020). Volunteering at the workplace under incomplete information: Teamsize does not matter. *MPI Collective Goods Discussion Paper*, (2020/4).

Hillenbrand, A. and Winter, F. (2018). Volunteering under population uncertainty. *Games and Economic Behavior*, 109:65–81.

Holt, C. A. (1986). Preference reversals and the independence axiom. *The American Economic Review*, 76(3):508–515.

Kopányi-Peuker, A. (2019). Yes, i'll do it: A large-scale experiment on the volunteer's dilemma. *Journal of Behavioral and Experimental Economics*, 80:211–218.

Krupka, E. L. and Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524.

Latané, B. and Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin*, 89(2):308.

Leo, G. (2017). Taking turns. *Games and Economic Behavior*, 102:525–547.

Maniaci, M. R. and Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality*, 48:61–83.

McKelvey, R. D. and Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38.

Otsubo, H. and Rapoport, A. (2008). Dynamic volunteer's dilemmas over a finite horizon: An experimental study. *Journal of Conflict Resolution*, 52(6):961–984.

Panchanathan, K., Frankenhuis, W. E., and Silk, J. B. (2013). The bystander effect in an n-person dictator game. *Organizational Behavior and Human Decision Processes*, 120(2):285–297.

Przepiorka, W. and Diekmann, A. (2018). Heterogeneous groups overcome the diffusion of responsibility problem in social norm enforcement. *PloS one*, 13(11):e0208129.

Ross, L. and Nisbett, R. E. (2011). *The person and the situation: Perspectives of social psychology*. Pinter and Martin Publishers.

Stahl, D. O. and Haruvy, E. (2006). Other-regarding preferences: Egalitarian warm glow, empathy, and group size. *Journal of Economic Behavior & Organization*, 61(1):20–41.

Tutic, A. (2014). Procedurally rational volunteers. *The Journal of Mathematical Sociology*, 38(3):219–232.

Weesie, J. (1993). Asymmetry and timing in the volunteer's dilemma. *The Journal of Conflict Resolution*, 37(3):569–590.

Weesie, J. (1994). Incomplete information and timing in the volunteer's dilemma a comparison of four models. *Journal of Conflict Resolution*, 38(3):557–585.

A Proofs and calculations

Proof of Proposition 1. By the definition of τ, $H_t < NH_T$ for $t > \tau$. Hence, bystanders do not help when $t > \tau$. For $t \leq \tau$, I first derive a symmetric equilibrium in behavioral strategies by using backward induction and then show that this is the unique symmetric subgame perfect equilibrium.

Let σ_τ be the probability that a bystander helps in period $t = \tau$ (by symmetry, this probability is the same for all bystanders). For a bystander to randomize, his utility of helping must be equal to his expected utility of not helping, such that

$$H_\tau = NH_T\left(1-\sigma_\tau\right)^{n-1} + NH_\tau\left[1-\left(1-\sigma_\tau\right)^{n-1}\right].$$

Rearranging the terms yields[20]

$$\sigma_\tau = 1 - \left(\frac{NH_\tau - H_\tau}{NH_\tau - NH_T}\right)^{\frac{1}{n-1}}.$$

Note that the bystander's expected payoff when $t = \tau$ is H_τ. Hence, for the bystander to be indifferent between helping and not helping in period $\tau - 1$, he takes into account that if no one helps in this period and the game continues, he will obtain (in expectation) H_τ. Thus, for him to randomize in period $\tau - 1$, it must hold that

$$H_{\tau-1} = H_\tau\left(1-\sigma_{\tau-1}\right)^{n-1} + NH_{\tau-1}\left[1-\left(1-\sigma_{\tau-1}\right)^{n-1}\right].$$

Then, rearranging,

$$\sigma_{\tau-1} = 1 - \left(\frac{NH_{\tau-1} - H_{\tau-1}}{NH_{\tau-1} - H_\tau}\right)^{\frac{1}{n-1}}.$$

The same procedure can be repeated backward until $t = 0$. Hence, the probability σ_t in equilibrium that a given bystander helps at period $t \in \{0, 1, \ldots, \tau - 1\}$ is

$$\sigma_t = 1 - \left(\frac{NH_t - H_t}{NH_t - H_{t+1}}\right)^{\frac{1}{n-1}}.$$

[20]Note that $\sigma_\tau = 0$ in the special case where $H_\tau = NH_T$.

To show that this is the unique symmetric subgame perfect equilibrium, note that there cannot exist an equilibrium in which all bystanders help at the same time since $NH_t > H_t$. There cannot exist an equilibrium in which bystanders help with probability $p_t \neq \sigma_t$ because this strategy would be dominated by either helping (if $p_t > \sigma_t$) or not helping (if $p_t < \sigma_t$). There cannot exist an equilibrium in which bystanders never help because $H_0 > NH_T$. Finally, there cannot exist an equilibrium in which for a single period $t < \tau$ bystanders do not help, since their expected utility in that period would then be $H_{t+1} < H_t$. □

The calculations of *Hypothesis 1* and *Hypothesis 2* are straightforward.

Calculation of Hypothesis 3. I first compute the mixed strategy equilibrium associated with the above equilibrium in behavioral strategies. The probability σ_t can be used to assign a probability to each of the possible pure strategies that a bystander can use.

Let s_t be the probability that a bystander helps at moment t given that he did not help earlier and conditional on no other bystander having helped before. Then, for $n \geq 2$,

$$s_t = \sigma_t \prod_{s=0}^{t-1} (1 - \sigma_s) .$$

Hence, the cumulative probability that a bystander has helped at period t is

$$S_t \equiv \sum_{z=0}^{t} s_z = \sum_{z=0}^{t} \sigma_z \prod_{s=0}^{z-1} (1 - \sigma_s) = 1 - \prod_{s=0}^{t} \left(\frac{NH_s - H_s}{NH_s - H_{s+1}} \right)^{\frac{1}{n-1}} .$$

Define $a_t \equiv \prod_{s=0}^{t} \frac{NH_s - H_s}{NH_s - H_{s+1}}$ and note that $a_t \in (0, 1)$. Then,

$$\frac{\partial S_t}{\partial n} = \frac{log\,(a_t) \cdot a_t^{\frac{1}{n-1}}}{(n-1)^2} < 0.$$

Furthermore, because $S_t = 1$ when $n = 1$ and $S_t < 1$ when $n \geq 2$ for every $t \in \{0, 1, \ldots, T - 1\}$, this implies that S_t decreases with n for every $n \geq 1$. □

Calculations of Hypothesis 4. Note that the probability S_{NH} that a bystander never helps is $S_{NH} = 1 - S_{T-1}$. Because S_t decreases with n for every $t \in \{0, 1, \ldots, T - 1\}$, it follows that S_{NH} increases with n. □

Calculations of Hypothesis 5. I first compute the cumulative probability that the victim is helped at some period and then show that it decreases with n. Let p_t be the probability of game termination at period $t \in \{0, \ldots, \tau - 1\}$. Note that for $n \geq 2$, the probability of no termination before t is $\prod_{s=0}^{t-1} (1 - \sigma_s)^n$. Then,

$$p_t = [1 - (1 - \sigma_t)^n] \prod_{s=0}^{t-1} (1 - \sigma_s)^n .$$

Therefore, the cumulative probability that the victim is helped at period t is

$$P_t \equiv \sum_{z=0}^{t} p_z = \sum_{z=0}^{t} (1 - (1 - \sigma_z)^n) \prod_{s=0}^{z-1} (1 - \sigma_s)^n =$$
$$= 1 - \prod_{s=0}^{t} \left(\frac{NH_s - H_s}{NH_s - H_{s+1}} \right)^{\frac{n}{n-1}} .$$

Define $a_t \equiv \prod_{s=0}^{t} \frac{NH_s - H_s}{NH_s - H_{s+1}}$ and note that $a_t \in (0, 1)$. Then,

$$\frac{\partial P_t}{\partial n} = \frac{log\,(a_t) \cdot a_t^{\frac{n}{n-1}}}{(n-1)^2} < 0.$$

Furthermore, because, $P_t = 1$ when $n = 1$ and $P_t < 1$ when $n \geq 2$ for every $t \in \{0, 1, \ldots, T - 1\}$, this implies that P_t decreases with n for every $n \geq 1$. □

B Two types of bystanders

This appendix consists of two parts. First, it shows that including *selfish bystanders* in the theoretical model preserves *Hypotheses 2*, *3*, and *4*. Then, it shows with a counterexample that *Hypothesis 5* no longer holds in this case.

The model uses the same assumptions as the model in Section 2, but it considers a population where there is a fraction $\gamma \in (0, 1]$ of bystanders for which $H_0 > NH_T$ (altruistic bystanders) and a fraction $1 - \gamma$ of bystanders for which $H_0 < NH_T$ (selfish bystanders). While altruistic bystanders would be willing to help if they knew that no one else would help, selfish bystanders are never willing to help.

Hypothesis 1 in Section 2 states that bystanders who are alone help immediately. In this model, this hypothesis would change to "bystanders *who help* do so immediately."

Hypothesis 2 predicts that bystander i, who is altruistic with probability γ and selfish with probability $1 - \gamma$, may help after some periods. Note that the complete information assumption in this context implies that bystander i is aware of how many other altruistic and selfish bystanders there are in his group.[21] This means that if he is an altruistic bystander, as long as there is another altruistic bystander in the group, he will use a mixed strategy. Because the main prediction of *Hypothesis 2* is that bystander i may spread his decision to help across time when groups are larger than one bystander, and because bystander i has a positive probability of facing another altruistic bystander, the prediction holds.

Hypotheses 3 and *4* predict that the cumulative probability with which a bystander helps at any period decreases with group size and that the probability that a bystander never helps increases with group size. To show that both hypotheses hold in this model, I will prove that the average probability that one bystander has helped at any period decreases with group size. Let bystander i be an altruistic bystander. Let $N \geq 0$ be the total number of other bystanders (a part of bystander i) in a group and n be the number of other altruistic bystanders in the same group (such that $N \geq n$). Because the only players who may help are the altruistic bystanders, then from *Hypothesis 3* the cumulative probability that bystander

[21] While this assumption may sound unrealistic when there are two types of bystanders, it makes it possible to derive analytical solutions without changing the results qualitatively (Bliss and Nalebuff 1984 and Weesie 1993; 1994 yield similar results in models with incomplete information and an infinite horizon).

i has helped at period t is

$$S_t \equiv 1 - \prod_{s=0}^{t} \left(\frac{NH_s - H_s}{NH_s - H_{s+1}} \right)^{\frac{1}{n-1}}.$$

Define $B_t(N)$ as the average cumulative probability that bystander j helps at period t or earlier when there are N other bystanders. Then, because bystander j has a γ probability of being altruistic,

$$B_t(0) = \gamma,$$

$$B_t(1) = \gamma\left((1-\gamma) + \gamma(1-a_t)\right),$$

$$B_t(2) = \gamma\left((1-\gamma)^2 + 2\gamma(1-\gamma)(1-a_t) + \gamma^2\left(1 - a_t^{\frac{1}{2}}\right)\right),$$

where $a_t \equiv \prod_{s=0}^{t} \left(\frac{NH_s - H_s}{NH_s - H_{s+1}} \right)$. In addition, for any N,

$$B_t(N) = \gamma\left(\sum_{n=1}^{N} \binom{N}{n} \gamma^n (1-\gamma)^{N-n} \left(1 - a_t^{\frac{1}{n}}\right) + (1-\gamma)^N\right),$$

and hence

$$\begin{aligned} &B_t(N-1) \\ &= \gamma\left(\sum_{n=1}^{N-1} \binom{N-1}{n} \gamma^n (1-\gamma)^{N-1-n} \left(1 - a_t^{\frac{1}{n}}\right) + (1-\gamma)^{N-1}\right). \end{aligned}$$

Showing that $B_t(N-1) - B_t(N) > 0$ completes the proof. Rewrite

$$\begin{aligned} &B_t(N) \\ &= \gamma\Bigg((1-\gamma)^{N-1}(1-a_t) + \sum_{n=2}^{N} \binom{N}{n} \gamma^n (1-\gamma)^{N-n} \left(1 - a_t^{\frac{1}{n}}\right) \\ &+ (1-\gamma)^N\Bigg). \end{aligned}$$

Then,

$$\begin{aligned}
&B_t(N-1) - B_t(N) \\
&= \gamma\Bigg(\left(\sum_{n=1}^{N-1} \binom{N-1}{n} \gamma^n (1-\gamma)^{N-1-n} \left(1 - a_t^{\frac{1}{n}}\right) + (1-\gamma)^{N-1} \right) \\
&- \left(\gamma (1-\gamma)^{N-1} (1-a_t) + \sum_{n=2}^{N} \binom{N}{n} \gamma^n (1-\gamma)^{N-n} \left(1 - a_t^{\frac{1}{n}}\right) \right. \\
&\left. + (1-\gamma)^N \right) \Bigg).
\end{aligned}$$

It is easy to see that

$$\begin{aligned}
&\sum_{n=1}^{N-1} \binom{N-1}{n} \gamma^n (1-\gamma)^{N-1-n} \left(1 - a_t^{\frac{1}{n}}\right) \\
&- \sum_{n=2}^{N} \binom{N}{n} \gamma^n (1-\gamma)^{N-n} \left(1 - a_t^{\frac{1}{n}}\right) \geq 0,
\end{aligned}$$

since both sums have $N-1$ components and for each one of these components (from smaller to larger n) the summation of both terms is positive.

This implies that if $(1-\gamma)^{N-1} - \gamma(1-\gamma)^{N-1}(1-a_t) - (1-\gamma)^N > 0$ then $B_t(N-1) - B_t(N) > 0$. Note that the LHS of the previous expression can be rewritten as $(1-\gamma)^{N-1}\gamma a_t$, which is always positive. Hence, it follows that $B_t(N-1) - B_t(N) > 0$, meaning that the cumulative probability that a given bystander has helped at any point in time decreases with group size. This implies that *Hypotheses 3* and *4* from Section 2 hold even when there are selfish bystanders.

As noted in Section 5, the same is not true for *Hypothesis 5.* This means that when some bystanders can be selfish, victims are not necessarily helped earlier nor are they more likely to be helped. To show that these hypotheses do not hold, take for example $\gamma = 0.5$. Then, when only one bystander can help the victim, 50% of the victims will be helped immediately, and 50% will never be helped. When two bystanders can help the victim, then with 25% probability both bystanders will be selfish and no one will help. With 50% probability one bystander will be selfish and one bystander will be altruistic, and therefore the altruistic bystander will help immediately. Finally, with 25% probability both bystanders will be altruistic and therefore help immediately with a given $a_0 > 0$ probability (from playing in mixed strategies). Therefore, in this case the victim is helped earlier and with higher probability when there are two bystanders than when there is one.

C Simulated quantitative predictions of the model

This appendix shows bystanders' predicted behavior according to the extended model of Appendix B. I assume the simplest utility function in which bystanders only care about their own and the victim's payoff in the same way. Hence, I assume that $H_t = 20 + (20 - t)$ and $NH_t = 24 + (20 - t)$. I further assume that 70% of the bystanders are altruistic and 30% are selfish.

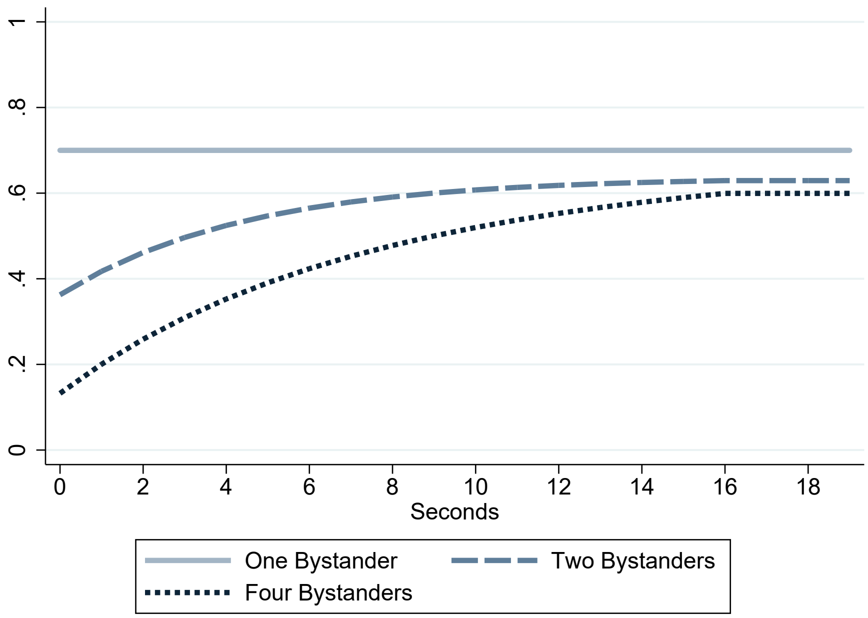

Figure A1. Simulated cumulative frequency of bystanders selecting to help at each second by group size.

Figure A.1. shows the predicted cumulative frequency of bystanders helping the victim at each second by group size. Note that the figure is fairly similar to the experimental results represented in Figure 2.

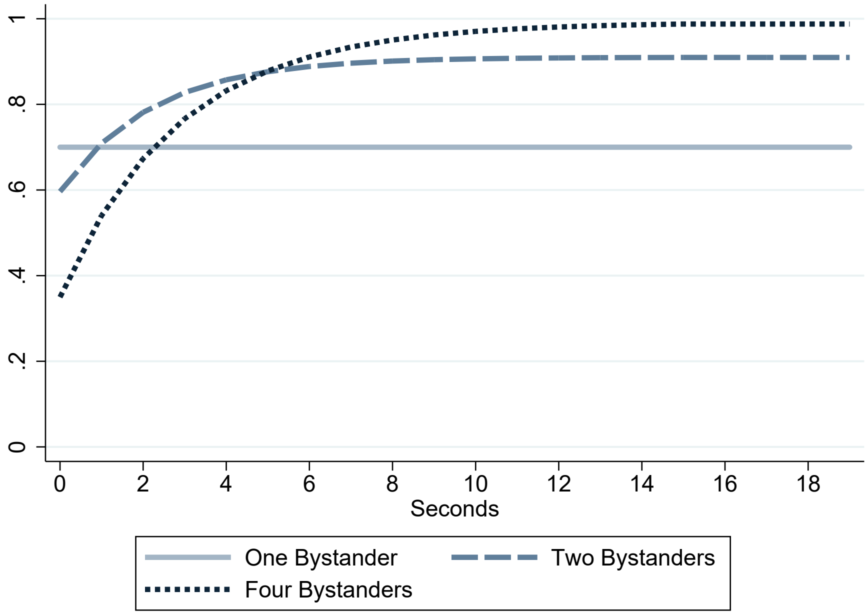

Figure A2. Simulated cumulative frequency of the victim being helped at each second by group size.

Figure A.2. shows the predicted cumulative frequency of victims being helped at each second by group size. Note that the figure is fairly similar to the experimental results represented in Figure 3.