SOCIAL IMAGE, OBSERVER IDENTITY, AND CROWDING UP*

Yamit Asulin[1,2], Yuval Heller[2], Nira Munichor[2], and Ro'i Zultan[1]

[1]Ben-Gurion University of the Negev
[2]Bar-Ilan University

January 14, 2024

Abstract

People behave more pro-socially when watched by others. Despite the abundance of theoretical and empirical research on social-image concerns, previous studies did not consider the observer's identity. To fill this gap, we develop a theoretical model incorporating social distance between the agent and the observer, and test the predictions of the model in a field experiment. 670 high-school students walked to generate donations for a public good. Participants were either unobserved, observed by a friend, or observed by an acquaintance. We also manipulated whether effort up to a certain threshold yielded a personal reward. Our results show that: (1) observability induces volunteers to exert more effort, (2) crowding up: a personal reward up to a threshold increases the share of volunteers who exert effort strictly above the threshold when the efforts are observed, and (3) among young adolescents, both effects are stronger when the observer is an acquaintance (rather than a friend).

Keywords: *Social distance, field experiment, crowding up, prosocial behavior.*
JEL codes: *C93,D64*

*We thank David Birke, Christine Exley, Sheen Levine, Friederike Reichel, Egon Tripodi, Erez Yoeli, and participants at the Helsinki GSE behavioral seminar, University of Birmingham, University of Amsterdam, Bar–Ilan University, University of Duisburg-Essen, the 7th Solomon Lew conference on behavioral economics in Tel Aviv, the Economic Science Association in Santa Barbara, Silvaplana Political Economy 2023, Switzerland, and SPUDM in Vienna for helpful comments and discussion. This study is part of Yamit Asulin's PhD. dissertation. All coauthors participated in the design of the experiment. YA conducted the experiment and collected the data. YH and RZ developed the theoretical model. YA and RZ analyzed the data and wrote the manuscript. All coauthors provided critical review and approved the manuscript. We thank the Eshkol Regional Council for funding the experiment. We are grateful for the Israeli Science Foundation for its financial support: Grant 448/22 for YH and YA, and grant 398/21 for RZ. **This article is dedicated to the participants in our experiment and their loved ones who tragically lost their lives in the October 7th, 2023, massacre in southern Israel.**

1. Introduction

People care about what others think of them. They consequently adjust their behavior when being observed to make a desirable impression on the observer. In particular, observation increases norm compliance. This phenomenon, known as social-image concern, affects a range of behaviors, including voting (DellaVigna et al., 2016; Funk, 2010; Perez-Truglia and Cruces, 2017; Rogers, Ternovski, and Yoeli, 2016), charitable giving (Alpizar, Carlsson, and Johansson-Stenman, 2008; Ariely, Bracha, and Meier, 2009; DellaVigna, List, and Malmendier, 2012), honesty (Bašić and Quercia, 2022), pro-environmental choices (Yoeli et al., 2013), ethical consumption (Friedrichsen and Engelmann, 2018), effort in school and the workplace (Ashraf, Bandiera, and Lee, 2014; Bursztyn, Egorov, and Jensen, 2019; Bursztyn and Jensen, 2015), and health-related behaviors (Karing, 2023).

Building on these effects, interventions designed to increase observability can promote pro-social behavior (Kraft-Todd et al., 2015). In designing such interventions, it is important to understand the moderating variables that shape responses to social-image concerns. Bursztyn and Jensen (2017) identified three such classes of moderators: the ability to infer the agent's type from her observed behavior, which primarily depends on observability; prevailing social norms, as reflected in the perceived desirability of being of a certain type; and the weight that the agent places on the opinions of others. Laboratory and field evidence support this analysis. Transparency and reduced anonymity increases contributions to public goods (Steiger and Zultan, 2014; Andreoni and Petrie, 2004), especially when the contributions are framed positively (Rege and Telle, 2004). The effect of observability depends on social norms (Graf et al., 2023). It is reduced and may even flip if the behavior is perceived to be socially *un*desirable (Ariely, Bracha, and Meier, 2009; Bursztyn, Egorov, and Jensen, 2019).

In this paper, we extend this analysis to introduce a novel moderator, namely the observer's identity, and in particular the social distance between the agent and the observer. Naturally, people care more about the opinions of friends than about the opinions of mere acquaintances.[1] People also conform more to peer pressure within one's immediate social circle or even those who are superficially similar to them (Bicchieri et al., 2022). Accordingly, DellaVigna et al. (2016) conjectured that social-image effects are stronger for observation by friends or family than observation by a stranger, writing that "we estimate the social-image utility when asked by a

[1] High social proximity should trigger reputation concerns, as it indicates a high likelihood of future interactions. This is consistent with Yoeli et al. (2013) finding a stronger effect of observability for home owners and people living in apartments, who have more direct and long-term relations with their neighbors.

stranger. If social-image concerns or lying costs when interacting with friends, family and colleagues are higher, our estimates are likely to be lower bounds of the social-image value of voting."[2] In contrast, we suggest that in certain situations, friends may have a weaker impact on behavior than mere acquaintances. The reasoning is straightforward within a signaling framework. An action may carry a strong influence on the perceptions of a casual acquaintance who lacks in-depth knowledge about the individual, but it may have a diminished impact on the opinions of a close friend who has intimate knowledge of the individual. This stands in contrast to other theories, such as social facilitation.[3]

While the existing economics literature on social-image concerns did not test the role of observer identity, several theoretical and empirical studies looked at the role of the observer's prior beliefs.[4] Adriani and Sonderegger (2019) showed theoretically that truncation of the prior type distribution reduces signaling incentives. We suggest that such a truncation can refer not only to the distribution of types in the population, as in Adriani and Sonderegger (2019), but also result from personal contact with the agent. Exley (2018) showed, theoretically and experimentally, that social-image concerns diminish if a history of volunteering is available to the observers. While the focus in that paper is on the prior exposure of the agent, we can construe social distance as differential prior exposure to different observers.

We build on a new clever method to provide a clean identification of social-image concerns developed by Birke (2023), who analyzed the effects of providing a personal reward for pro-social activities conditional on hitting a certain threshold. Sans social-image concerns, such a reward should lead to bunching at the threshold. Birke (2023) showed, theoretically and in an online experiment, that the reward undermines the signaling value of the activity. Consequently, the reward leads to anti-bunching under observation: pro-social individuals perform above and beyond the level required to

[2]Similarly, Meyer and Tripodi (2021) attribute their results to higher sensitivity to observation by friends or family. In their field experiment, people accompanied friends and family—but not people surrounded by strangers—responded more positively to verbal (and therefore public) offers to sign up for a blood drive compared to offers made privately on a tablet screen. However, the observers' identity was endogenous and not part of the experimental design, exposing the results to selection problems. Moreover, a visibility measure indicated that friends and family were more likely to overhear the public offer and the following response than strangers, as reported in an early version of the paper. Controlling for this measure, however, did not explain away audience effect.

[3]Social facilitation theories assert that the presence of (or being observed by) others increases arousal (Bond and Titus, 1983; Zajonc, 1965). Studies comparing responses of presence of friends and strangers found higher arousal and activity in response to friends (Clendenen, Herman, and Polivy, 1994; Wagner and Smith, 1991).

[4]In the psychological literature, McKelvey and Kerr (1988) found that people conform in the Asch (2016) paradigm more to a group of strangers than to a group of friends. Tice et al. (1995) found that people present themselves more modestly to friends than to strangers. The authors offer several possible explanations, including the need for making a good first impression on a stranger or the ability of friends to dispute excessively positive claims. Li et al. (2022) found that observability increases giving to acquaintances and strangers but not to friends, although this may be explained by a ceiling effect.

obtain the reward in order to separate from those motivated by the reward. We term this phenomenon *Crowding up*.[5]

We model social distance and crowding up within a simplified discrete framework based on Bénabou and Tirole (2006) with three possible actions: no, low, and high contribution. We consider three types, each type preferring a different level of contribution in the baseline setup without observation. With observation, the two lower types increase their equilibrium contribution by one level with positive probability. We refer to this prediction as the *observability effect*. A reward for any contribution, whether low or high, increases the probability that a low type chooses high contribution *only* under observation. We term this prediction the *crowding-up effect*.

We introduce social distance with two assumptions. First, the agent assigns a higher weight to a friend's beliefs about her type than to an acquaintance's beliefs. Second, we assume that a friend observer, unlike an acquaintance, knows the agent type with positive probability. The interplay of these two assumptions, expressed formally as the product of the weight factor and the probability that a friend knows the agent's type, determines whether the observation and crowding-up effects increase or decrease when the observer is a friend. Thus, the *observer identity* effect predicts that both the observability and crowding-up effects may be stronger when an acquaintance rather than a friend is observing.

We tested these three effects in a field experiment with 670 secondary school students. We teamed up with a regional council in the south of Israel as part of a social sports project. For each ten steps that a student walked within a one-hour period, the regional council pledged one New Israeli Shekel towards a new community sports center. We manipulated the personal reward and observability in a 2×3 between-subjects design resulting in six treatment groups. For participants in the *Reward* (but not in the *No reward*) treatments, participation up to 3,000 steps counted towards mandatory volunteering hours. Crossed with the reward manipulation, we manipulated whether the participant's performance (and reward, if relevant) remained confidential, were conveyed to a friend of the participant, or conveyed to another random student.

The results confirm our predictions. Without reward, observation increases the mean number of steps taken by 19.2% and the median number by 31.9%. The reward has no significant effect without observation, with 47.8% of participants who received the reward walking more than 3,000 steps compared to 51.4% of participants randomly assigned not to receive the reward. With observation, the reward crowds up performance, increasing the share of participants surpassing the threshold from 65.2% without reward to 87.2% with reward.

[5] This is a special form of crowding in. It encompasses discrete activity levels, whereas *anti-bunching* describes continuous contributions.

Distinguishing between observation by a friend and an acquaintance, we find that observation by an acquaintance increases the number of steps taken by 23.7% compared to 15.3% when a friend is observing, however, this difference is not statistically significant. A post-hoc analysis reveals that this effect varies dramatically by age. Younger participants, up to the tenth grade, increase their efforts more if an acquaintance is observing (34.1% vs. 11.6%). In contrast, older participants respond stronger to a friend observing and almost not at all to observation by an acquaintance (23.5% vs. 3.5%). The crowding-up effect is also stronger for observation by an acquaintance versus a friend, with the reward increasing the share of participants who surpass the threshold by 30.2% when a friend is observing and by 39.6% when an acquaintance is observing. Here too, observation by an acquaintance has a stronger effect on the younger participants while the share of older participants exceeding the threshold does not differ by observer identity.

This paper makes four contributions. First, to the best of our knowledge, this is the first field experiment that experimentally varies the identity of the observer and tests the resulting effect on the observed agent's behavior. Previous studies manipulated whether participants were observed or not. DellaVigna et al. (2016) found that exposure to a stranger surveyor was sufficient to increase willingness to vote. In other studies, the observers were members of the relevant community (Ariely, Bracha, and Meier, 2009; Bursztyn, Egorov, and Jensen, 2019; Bursztyn and Jensen, 2015; Karing, 2023). None of these manipulated observer identity within the same study.

Second, we extend the theoretical analysis of social-image concerns to include different observers. Theoretical models typically assume a commonly known prior distribution of types. Some treatments considered the possibility of different agents being associated with different prior distributions due to different peer groups (Adriani and Sonderegger, 2019) or prior history (Exley, 2018). We show that a similar analysis can encompass different observers.

Third, we are the first to validate the procedure suggested by Birke (2023) to identify social-image concerns through crowding up in field settings. The role of external incentives in social image was studied in the context of crowding out: providing incentives reduces the effect of being observed (Ariely, Bracha, and Meier, 2009). In Birke's (2023) online experiment, incentives at a given level interact with observability to increase effort. We show that the effect can be extended to field settings.

Fourth, our results contribute to the literature on interventions designed to promote normative behaviors through increased observability. For example, interventions leveraging exposure on social media (e.g., Cameron et al., 2013) can improve their efficacy by targeting specific audiences. Moreover, in our sample of adolescents, age turns out to be an important moderator. More broadly, our results indicate that different types of observers may have stronger effects depending on the characteristics of the agent.

The remainder of this paper is organized as follows. Section 2 develops the theoretical framework. Section 3 describes the field experiment. Section 4 reports the results of the experiment. Section 5 discusses the results and concludes. Supplementary materials, including replication materials, can be found at the OSF project (Asulin et al., 2023).

2. Theoretical framework

In this section we present a special case of the model of Bénabou and Tirole (2006) to generate testable predictions for crowding up and the role of observer identity. To avoid complexities that do not bear on the model predictions, we keep the setup minimal, with three possible actions and corresponding types.[6]

2.1. Model

Types, Actions and Basic Payoffs An agent has to choose her level of effort in a socially desirable activity. There are three possible actions (interpreted as levels of effort): $A = \{a_0, a_1, a_2\}$, where action a_j is interpreted as exerting effort level j. The agent has one of three possible types $\theta_0, \theta_1, \theta_2 \in \Theta$, where type θ_i has prior probability $q_i \geq 0$ (with $q_0 + q_1 + q_2 = 1$) . The agent knows her own type. The net utility of type i from playing action a_j is u_{ij} (which reflects the agent's benefit from the socially desirable activity and the agent's cost of effort). We assume that the optimal action of type i is a_i (i.e., $u_{ii} > u_{ij}$ for each $i \neq j$).

Let $0.5 < c_1 := u_{00} - u_{01} < 1$ denote the disutility of a player of type θ_0 from playing action a_1 instead of her favorite action (a_0). Similarly, let $0.5 < c_2 := u_{11} - u_{12} < 1$ denote the disutility of player of type θ_1 from playing action a_2 instead of her favorite action (a_1). We further assume that the disutility of a player of type θ_0 from playing action a_2 is sufficiently large, namely $u_{02} < u_{01} - 2$.

In the baseline setup, the payoff of an agent of type θ_i from playing action a_j is u_{ij}. In a setup with a *reward*, the agent gets an additional payoff of $b > 0$ if her action is either a_1 or a_2 (interpreted as a personal bonus payment for effort surpassing a threshold). Thus, the total payoff of an agent of type θ_i from playing action a_j is $u_{ij} + b$ if $j \geq 1$ and u_{i0} if $j = 0$.

[6] Three actions is the minimal number required to capture crowding up, which involves three levels of contribution: the threshold level required to obtain the personal reward, below-threshold level, and strictly-above threshold.

Strategies and Reputation Payoffs A *strategy* is a function $\sigma : \Theta \rightarrow \Delta(A)$ that specifies the distribution of actions of the agent as a function of her type. Let $\sigma_{\theta_i}(a_j)$ denote the probability in which an agent of type θ_i plays action a_j when following strategy σ. Let $P_\sigma(\theta_i|a_j)$ denote the posterior probability that an agent has type θ_i, conditional on the agent following strategy σ and on playing action a_j. Let $E_\sigma(\theta|a_j) := \sum_i i \cdot P_\sigma(\theta_i|a_j)$ denote the expected type of an agent, conditional on the agent following strategy σ and on playing action a_j.

With a very small probability $0 < \epsilon << 1$ the agent does not care about reputation, and her payoff with observability is the same as without observability. With the remaining probability of $1-\epsilon$, the agent obtains additional payoff based on the observer's assessment of the agent's type. Specifically, when agent's action is observed by an acquaintance (who does not know the agent's type), the agent obtains $E_\sigma(\theta|a_j)$ as an additional payoff (henceforth, *reputation payoff*).

Finally, when the agent's action is observed by a friend, we assume that there is a probability $\mu \in [0,1]$ that the friend knows the agent's type (in which case the friend's assessment of the agent's type is not affected by the agent's action).[7] The agent does not know whether the friend knows the agent's type or not. With probability $1-\epsilon$ the agent has an additional payoff that is equal to $d > 0$ times the expected type that the friend's belief assigns to the agent conditional on the agent's action. Thus, in this case, the total payoff of an agent of type θ_i from playing action a_j is $u_{ij} + d \cdot \left(\mu \cdot \theta_i + (1-\mu) \cdot E_\sigma(\theta|a_j)\right)$. Thus, the parameters μ and d capture the two dimensions of social distance: how much the agent cares about her reputation and the observer's familiarity with the agent, respectively. We assume that $c_1, c_2 < d \cdot (1-\mu)$. The payoffs in all $6 = 2 \times 3$ cases are summarized in Table 1.

Equilibrium and Additional Assumptions Let $\pi_\sigma(\theta_i)$ denote the expected payoff of an agent of type θ_i who follows strategy σ (and outside observers have the correct equilbirium belief of the agent following strategy σ). Let $\pi_\sigma(a_j|\theta_i)$ denote the expected payoff of an agent of type θ_j who plays action a_i, while observers still believe the player to follow strategy σ. Strategy σ is a (Bayes-Nash) equilibrium if $\pi_\sigma(\theta_i) \geq \pi_\sigma(a_j|\theta_i)$ for each type θ_i and each action a_j.

Let $E_\sigma(a_i)$ denote the expected level of effort of an agent who plays strategy σ:

$$E_\sigma(a_i) = \sum_j q_j \cdot \sigma(\theta_j)(a_1) + 2 \cdot \sum_j q_j \cdot \sigma(\theta_j)(a_2).$$

We show later that the environment admits a unique equilibrium. For tractability, we make a few assumptions on the parameters that imply that this equilibrium

[7] Our results hold if we assume that an acquaintance knows the agent's type with a positive probability smaller than μ.

Table 1: Payoffs in the Various Setups

		Payoff of agent of type θ_i playing action a_j	
Reward	Observability	a_0	a_1 or a_2
No	No	u_{ij}	
No	Acquaintance	$u_{ij} + E\left(\theta \vert a_j\right)$	
No	Friend	$u_{ij} + d\mu\theta_i + (1-\mu)\, E\left(\theta \vert a_j\right)$	
Yes	No	u_{i0}	$u_{ij} + b$
Yes	Acquaintance	$u_{i0} + E\left(\theta \vert a_0\right)$	$u_{ij} + E\left(\theta \vert a_j\right) + b$
Yes	Friend	$u_{i0} + \mu\theta_i +$ $d\,(1-\mu)\, E\left(\theta \vert a_0\right)$	$u_{ij} + \mu\theta_i +$ $d\,(1-\mu)\, E\left(\theta \vert a_j\right) + b$

Notes: No observability includes cases where the agent does not care about reputation.

is "interior" in the setup with observability and without a reward in the sense that an agent of type θ_0 or θ_1 plays a mixed action (type θ_i mixes between action a_i and action a_{i+1}). All of our results can be extended to a setup without these assumptions, but this will make our notation and statement of result somewhat more cumbersome, as they will have to deal also with the cases of "corner" equilibria in which an agent of type θ_0 or θ_1 plays a pure action in the setups with observability. The assumptions are: $q_1 \leq q_0$ and $\frac{q_2(2v-c_1-c_2)}{q_1(c_2+c_1-v)} < 1$.

2.2. Results and testable implications

We begin with a straightforward characterization of the unique equilibrium without observability: the two higher types play their respective actions, while the low type θ_0 plays a_0 if the reward is low and a_1 if the reward is high.

Claim 1. *Without observability, the equilibrium strategy is as follows: agent of type θ_1 always plays a_1, agent of type θ_2 always plays a_2, and agent of type θ_0 always plays a_0 if $b < c_1$, always plays a_1 if $b > c_1$, and she mixes between a_0 and a_1 if $b = c_1$.*

Proof. The payoff structure immediately implies that each agent of type $\theta_i \neq \theta_0$ maximizes her payoff by playing a_i in the setups without observability, and that an agent of type θ_0 maximizes her payoff by playing a_0 if $b < c_0$ and by playing a_1 if $b > c_0$. □

Our first result shows the impact of observability in our setup is inducing the agent to play with positive probability an action that is one level above her type. Specifically, we show that the environment admits a unique equilibrium, in which type θ_0 either plays a_0 or a_1, type θ_1 either plays a_1 or a_2, and type θ_2 plays a_2.

Table 2: Example ($q_0 = q_1 = 40\%$, $q_2 = 20\%$, $b = 0.1$, $c_1 = c_2 = 0.8$, $d \cdot (1-\mu) = 0.95$)

#	Reward	Observation	$\sigma(a_1 \mid \theta_0)$	$\sigma(a_2 \mid \theta_1)$
1	No	No	0%	0%
2	No	Acquaintance	17%	33%
3	No	Friend	14%	23%
4	Yes	No	0%	0%
5	Yes	Acquaintance	25%	50%
6	Yes	Friend	27%	36%
	All cases		$\sigma(a_0 \mid \theta_0) = 1 - \sigma(a_1 \mid \theta_0)$	$\sigma(a_1 \mid \theta_1) = 1 - \sigma(a_2 \mid \theta_2)$

Proposition 1. *The environment admits a unique equilibrium σ. In this equilibrium any type θ_i either plays action a_i or a_{i+1}.*

The proof of Proposition 1 (which appears in Appendix A) fully characterizes the unique equilibrium in all six cases. This full characterization implies various interesting corollaries, that informs our hypotheses for the experiment. Table 2 demonstrates this characterization in an example in which $q_0 = q_1 = 40\%$, $q_2 = 20\%$, $b = 0.1$, $c_1 = c_2 = 0.8$, and $d\,(1-\mu) = 0.95$.

Our first corollary shows that observability induces a positive probability of an agent of type $\theta_i \neq \theta_2$ playing an effort level above her type.

Corollary 1. *Consider the setup with no reward. Adding observability (either by a friend or by an acquaintance) induces a positive probability for an agent of type θ_0 (resp., θ_1) playing action a_1 (resp., a_2).*

This implies the first testable prediction of our model, the obervability effect:

Prediction 1 (Observability)**.** *Observability increases voluntary effort in socially desirable activities.*

Next, we show that adding a reward to a setup with observability increases the probability that the agent exerts effort above the reward's threshold (the "crowding up" effect). That is, the reward, which is given up to anyone exerting at-least effort a_1, induces a higher probability of the agent exerting effort a_2. This occurs because the personal reward for playing action a_1 deteriorates the reputation of being observed playing a_1, and this discourages an agent of type θ_1 from playing action a_1. The fact that there is no additional reward for effort above a_1, allows the agent to separate herself from the agents of type θ_0 who play a_1 by exerting effort a_2 with a higher probability. We term this phenomenon *crowding up*.

Corollary 2. *Consider a setup with observability (either by an acquaintance or by a friend). Adding a reward increases the (crowding up) probability in which agent of type* θ_1 *playing action* a_2*.*

This induces our second prediction: the crowding up effect:

Prediction 2 (Crowding up)**.** *A personal reward up to a threshold increases the share of volunteers that exert efforts strictly above the threshold when the volunteers efforts are observed (and this does not happen when the efforts are not observed).*

For our final result, we assume that $d \cdot (1 - \mu) < 1$, which implies that the reputation effects are stronger for an observing acquaintance than to an observing friend. That is, although an agent may care more about the friend's opinion (i.e., having $d > 1$), the fact that the friend's opinion has a sufficiently high positive probability μ to be independent of the agent's action, implies that the indirect effect of the observed action on the agent's utility (through its impact on the observer's assessment of the agent's type) is larger when being observed by an acquaintance.

Corollary 3. *Assume* $d \cdot (1 - \mu) < 1$*. Changing the observability from a friend to an acquaintance:*

1. *increases the expected equilibrium level of effort* $E_\sigma(a_i)$ *in the setup without a reward, and*
2. *increases the (crowding up) probability of an agent of type* θ_1 *playing action* a_2 *in the setup with a reward.*

The opposite holds if $d \cdot (1 - \mu) < 1$ *(i.e., one should replace "increases" with "decreases" the corollary's statement).*

This induces our last prediction:

Prediction 3 (Observer identity)**.** *In some cases (where* $d \cdot (1 - \mu) < 1$*), observability by an acquaintance increases both the observability effect and the crowding up effect:*

1. *Without a personal reward: Observability by an acquaintance increases the mean effort level of volunteers in socially desirable activities relative to observability by a friend.*

2. *With a personal reward: Observability by an acquaintance increases the share of volunteers exerting efforts above the reward's threshold relative to observability by a friend.*

2.3. Discussion of assumptions

Our assumption that there is a small probability in which the agent does not care about her reputation does allow us to achieve equilibrium uniqueness. Specifically, if all agents care about reputation, then we have an additional equilibrium σ' in which an agent of type θ_0 plays a_0, and an agent of either type θ_1 or θ_2 plays a_2, and the observer has the off-the-equilibrium-path belief that if he observes action a_1 (which is never observed on the equilibrium path), then the agent's type is θ_0. We think that this equilibrium σ' is an implausible prediction for long-run behavior in populations who play this game. This is so because starting from σ' any small deviation of a few agents of type θ_1 who play a_1,would induce observers to the correct belief of agents playing action a_1 having type θ_1, which, in turn, would induce more agents of types θ_0 and θ_1 to play a_1 until converging to the equilibrium σ, which is characterized in Proposition 1.

3. Field experiment

We conducted a field experiment testing the predictions of the theoretical model (we state the operational hypotheses at the end of this section). We teamed up with the Eshkol Regional council in the south of Israel as part of a "Social Sports Project" initiative. As part of the project, 670 students from grades eight to twelves (ages 13–18) from the regional secondary school raised money for a new sports center by "walking for health." Each student walked for one hour with a dedicated smartphone application counting their steps. The regional council pledged to allocate 0.1 NIS (New Israeli Shekel) for each step walked within the hour to the new center, and up to a maximum amount of 300,000 NIS (approximately $94,000$ USD at the time of the experiment). We manipulated the personal reward and observer identity in a 2×3 between-subjects design. The experiment received ethical approval from the review board at Bar-Ilan University.

3.1. Experimental Design

The experiment included six distinct experimental groups resulting from crossing two levels of reward (*Reward* and *No reward*) with three levels of observation (*Anonymous*, *Friend*, and *Acquaintance*). The allocation to groups was random within each grade year and determined once the student activated the application. Table 3 presents the number of participants in each group by grade year.

Of the 670 participants, 225 received a personal reward for walking. The *Reward* treatments relied on the requirement for volunteering activities as part of the national conditions for matriculation. High-school students in Israel are required to complete 90 volunteering hours including 30 hours of group activities in the tenth grade, 60 hours in the eleventh grade, and 30 hours in the twelfth grade.[8] Because

[8]Volunteering encompasses diverse activities including, e.g., agriculture work, environmental

Table 3: Experimental design

	No Reward					
Grades	8th	9th	10th	11th	12th	Total
Anonymous	37	37	24	24	24	**146**
Friend	43	43	25	25	25	**161**
Acquaintance	37	38	21	21	21	**138**
	Reward					
Grades			10th	11th	12th	Total
Anonymous			23	23	23	**69**
Friend			26	26	26	**78**
Acquaintance			26	26	26	**78**

Notes: The table shows the number of participants in each cell. Allocation to treatments was randomized within each grade year.

this requirement holds for tenth-grade students and older, all of the participants in the eighth and ninth grades were allocated to the *No reward* groups, whereas older participants were randomly allocated between the *Reward* and *No reward* groups. For participants in the *Reward* groups, each 500 steps and up to 3,000 steps counted as 30 minutes of volunteering. The crowding-up effect predicts that (only) under observation, this reward increases the share of participants surpassing the threshold of 3,000 required to obtain the maximal reward. Eventually, all of the participants received credit for one hour of volunteering for their participation regardless of their performance. The participants were not aware of this until after the experiment, nor were the participants in the *No reward* groups aware that their peers received performance-based credit.

In the *Friend* and *Acquaintance* observer treatments, the application informed another student of the performance of the participant, including the accumulated volunteering hours if relevant. In the *Friend* treatments, this other student was a friend whose name and number the participant had provided in the preceding week while installing the application. In the *Acquaintance* treatments, the information was disclosed to a randomly selected student in the school. In the *Anonymous* treatments, all information remained confidential.

3.2. Experimental Procedure

We developed a smartphone application designed for Android and iPhone users to manage the experiment, send messages to observers, and record the data. The application was available for free download from the respective application stores. Ap-

work, and social work. See *Ministry of Education* (2014) for official information in Hebrew.

proximately a week before the experiment, student council representatives visited classrooms to explain how participants' actions would generate funds for a school project. The representatives ensured that the students downloaded the application. Each student received a unique code with which he or she could register. During registration, students entered demographic information including gender and grade, their contact details, and the name and number of a close friend. The application used the contact details and friend's number to send messages with participants' performance in the friend and acquaintance treatments but did not record them, or any identifying information, in the data.[9]

The experiment itself took place over a 24-hour period during the weekend. The application was accessible for activation only within this time window. Once activated, the application randomly assigned the participant to a treatment group, as described above, and informed the participant of the details specific to the treatment condition. The application remained activated for sixty minutes unless the participant chose to stop it earlier. During the walk, the screen showed the remaining time, number of steps walked, accumulated donation amount, and, in the *Reward* groups, the accumulated reward and steps remaining to the next reward level. At the end of the experiment, the application sent text messages notifying observers (friends or random participants, depending on the treatment) of the participant's donation amount and personal reward. After the end of the experiment, the application removed all identifying information to comply with ethical requirements, retaining only the treatment, step count, gender, age and grade year for each participant.

Note that the features of the experiment reduce concerns about leakage of information between different treatments. First, we conducted the experiment over the weekend. The school caters to 32 different localities ("Kibbutzim" and "Moshavim") spread across the region, with an average of fewer than 500 residents per locality. Consequently, only a few students of the same age group live in the same locality and are likely to be together during the weekend. Second, the treatment variation was kept confidential and was revealed to each participant only as they started walking.

3.3. Hypotheses

Our hypotheses operationalize the theoretical predictions in the context of the experiment. The first hypothesis corresponds to Prediction 1.

[9] The registration also included a consent form and additional questions: "Have you volunteered for an association before?" and "Do you have any special hobbies?" We added these questions to avoid drawing special attention to the question about a friend.

Hypothesis 1 (Observability)**.** *In the* No reward *treatment, the step count is higher in the observed groups compared to the anonymous group.*

Prediction 2 states that a reward increases the probability of surpassing the threshold required to obtain the reward. In the experiment, there are different levels of reward, and the prediction refers to the highest level above which there is no additional reward.

Hypothesis 2 (Crowding up)**.** *A reward increases the share of participants walking more than* 3,000 *steps in the observed groups.*

The theoretical analysis assumes that each individual chooses an exact effort level. The experimental setting, in contrast, may involve noise considerations not accounted for in the theory. For example, assume that the reward incentivizes some participants to walk 3,000 steps to obtain the full reward even without observation. These participants may have some time left or some distance to walk back home, leading to above 3,000 step counts. To account for such issues, we first include robustness checks extending the threshold to above 3,000 steps. In addition, we state the following stricter version of Hypothesis 2, using the anonymous groups to control for any effects of the reward other than crowding up.

Hypothesis 2′ (Crowding up)**.** *The effect of a reward on the share of participants walking more than* 3,000 *steps is stronger in the observed groups.*

After establishing the observability and crowding-up effects in general, we compare both effects between friend and acquaintance following Prediction 3.

Hypothesis 3a (Observer identity and observability)**.** *The mean step count is higher when observed by an acquaintance than when observed by a friend.*

Hypothesis 3b (Observer identity and crowding up)**.** *The share of participants walking more than* 3,000 *steps is higher when observed by an acquaintance than when observed by a friend.*

4. Results

Overall, the average step count was 3,448 steps, with a standard deviation of 1,487. Of the 670 participants, 430 (64%) walked more than 3,000 steps (the number of steps required to obtain the maximal reward, when applicable). Table 4 presents descriptive statistics by treatment group. The comparative statics are in line with our hypotheses. The step count without reward is highest for acquaintance observation and lowest in the anonymous group. The reward has close to no effect on the share

Table 4: Descriptive Statistics

	Reward groups		
	Anonymous	Friend	Acquaintance
Average step count	2,836	3,761	3,943
	(1,171)	(1,128)	(804)
Share > 3,000	48%	83%	91%
N	69	78	78
	No reward groups		
	Anonymous	Friend	Acquaintance
Average step count	3,013	3,475	3,726
	(1,458)	(1,738)	(1,610)
Share > 3,000	47%	64%	65%
N	146	161	138

Notes: Standard deviations in parentheses. The *Reward* groups include only tenth-grade students and older.

of participants walking more than 3,000 steps in the anonymous treatment, and has the strongest effect in the acquaintance treatment. Note, however, that the *No reward* groups include the younger students that were not eligible for rewards and thus were not included in the Reward groups. Hence, the net effect of reward is not estimable from the table. In the following, we address the different hypotheses in order.

4.1. Observability

Figure 1 presents the step-count distributions in the anonymous and observed (by either friend or acquaintance) groups without rewards. Observed participants walk, on average, $3,591$ steps compared to $3,013$ steps walked by anonymous participants (Mann-Whitney $z = 3.69, p < .001$). Column (1) of Table 5 reports a regression of the step count on observation confirming this conclusion.

Result 1. *Observed participants walk more than anonymous participants, confirming Hypothesis 1.*

4.2. Crowding Up

We restrict the analyses in this section to 10th grade students and older, who were eligible for the reward, to maintain comparability of those who did and did not receive the reward. Panels A and B in Figure 2 show the step-count distributions for observed participants with and without a reward. Consistent with crowding up, the reward increases the share of observed participants walking more than 3,000 steps from 65.2%

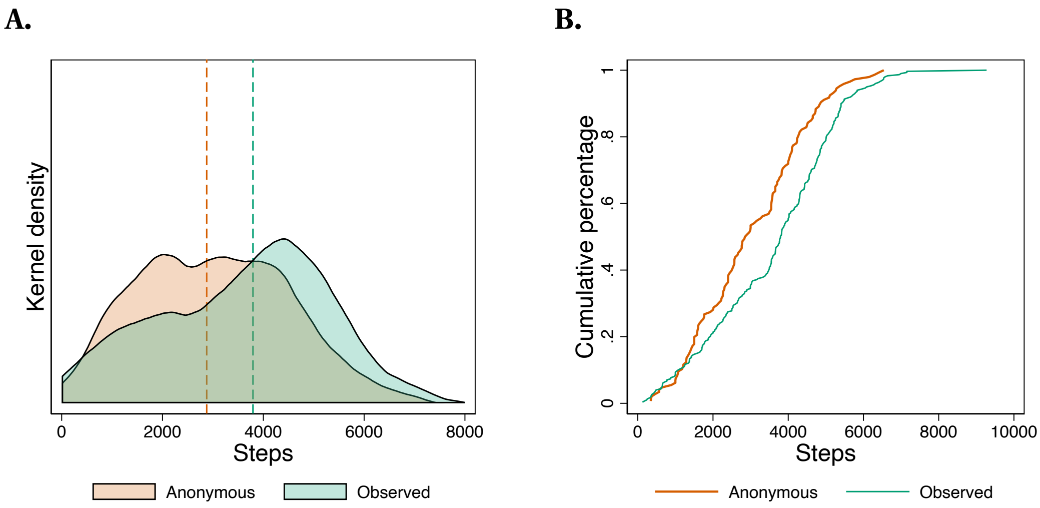

Figure 1: Step-count distributions without reward.

Notes: Kernel density function (**A**) and cumulative distribution function (**B**) of the step counts. *Observed* aggregates the *Friend* and *Acquaintance* groups. The vertical dashed lines in Panel **A** mark the medians of the groups.

Table 5: Regressions on step counts

	(1)	(2)	(3)	(4)
Observed	578***			
	(3.79)			
Anonymous		-462***	-471***	-206
		(-2.61)	(-2.67)	(-0.73)
Acquaintance		251	253	884***
		(1.36)	(1.35)	(2.97)
Grade			122**	292***
			(2.36)	(3.12)
Anonymous × Grade				-160
				(-1.30)
Acquaintance × Grade				-382***
				(-2.94)
Constant	3013***	3475***	3272***	2989***
	(25.04)	(26.00)	(20.38)	(13.66)
Observations	445	445	445	445

Notes: OLS regressions with bootstrap standard errors stratified within grade years. *Grade* is normalized so that the 8th grade is the baseline. t-statistics in parentheses. $^{*}\ p < .10$, $^{**}\ p < .05$, $^{***}\ p < .01$.

to 87.2% (Fisher's exact $p < .001$). This increase is clearly due to observation, as the reward had, if at all, the opposite effect on anonymous participants, decreasing the share of participants walking more than 3,000 from 51.4% to 47.8% (Fisher's exact $p = .737$, see Panels C and D in Figure 2).

Column (1) of Table 6 presents the results of a logistic regression with a dummy for over 3,000 steps as a dependent variable and (no) reward interacted with (no) observation as independent variables. Analyzing the impact of receiving a reward instead of its absence, the regression results indicate that reward has a positive and significant effect under observation ($z = 4.47, p < .001$) and a negative and non-significant effect in the anonymous treatments ($z = -0.40, p = .686$), confirming Hypothesis 2. Figure 3 presents these results graphically.

Consistent with Hypothesis $2'$, the interaction of reward and observation is also highly significant ($z = 3.22, p = .001$), indicating that the effect of the reward is indeed due to crowding up and not to any other mechanism that should also be relevant in the anonymous treatments. Furthermore, the results qualitatively hold if we set the higher criterion of 3,300 steps, as in Column (2) of the table, with a positive significant effect for reward under observation ($z = 2.86, p = .004$) and a non-significant negative effect in the anonymous treatments ($z = -1.09, p = .276$).

Result 2. *A reward crowds up efforts to surpass the reward threshold only under observation, confirming Hypotheses 2 and $2'$.*

4.3. Observer identity

We first test the moderating effect of observer identity on the observability effect, as stated in Hypothesis 3a. Figure 4 presents the distribution of steps by observer identity. Column (2) of Table 5 reports the results of regressing the step count on observer identity with observation by friend as the baseline. Those observed by a friend walked, on average, 462 steps more than the anonymous participants ($t = 2.61, p = .009$). Observation by an acquaintance increases the step count by 713 steps, 54.3% more than for a friend. The difference, however, is not statistically significant ($t = 1.36, p = .174$).

Result 3. *The step count is higher for observation by an acquaintance compared to observation by a friend. The difference is, however, not statistically significant. Thus, the results do not confirm Hypothesis 3a.*

After failing to find an overall significant effect for observer identity, we turn to study the observer identity effect by grade year. Figure 5 shows the step counts by observer identity for the different grade years. Table 7 reports treatment comparisons for observer identity by grade year. These comparisons are based on an OLS

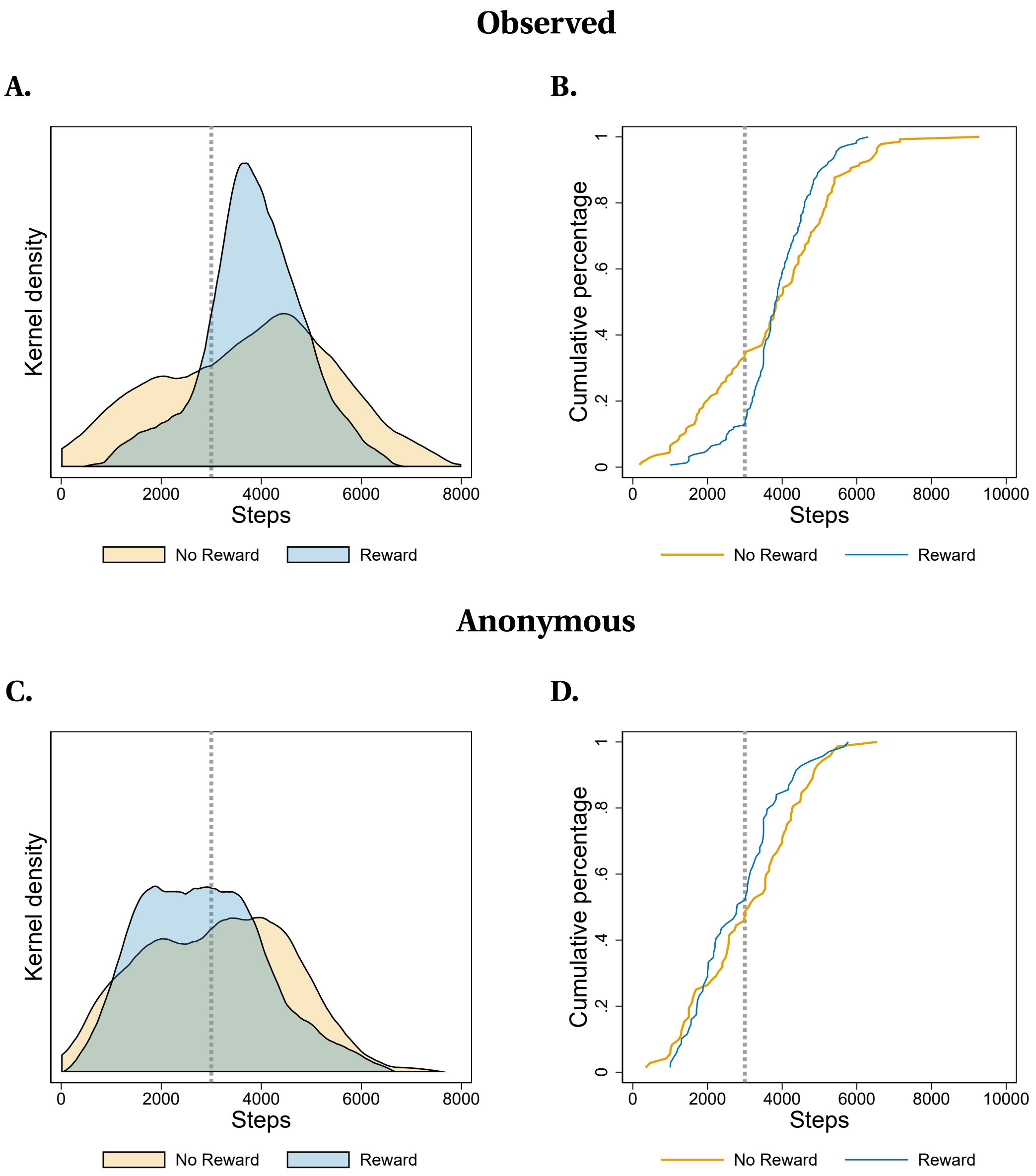

Figure 2: Step-count distributions, 10th grade and older.

Notes: Kernel density function (**A**) and cumulative distribution function (**B**) of the step counts for observed participants and kernel density function (**C**) and cumulative distribution function (**D**) of the step counts for anonymous participants. The vertical dotted line mark the 3,000-steps threshold.

Table 6: Regressions on crowding up

	(1) $> 3,000$	(2) $> 3,300$	(3) $> 3,000$
Anonymous	-2.004***	-1.796***	
	(-5.81)	(-5.41)	
No reward	-1.288***	-0.758***	0.143
	(-4.47)	(-2.82)	(0.42)
Anonymous × No reward	1.431***	1.144**	
	(3.22)	(2.54)	
Friend			1.696***
			(4.34)
Acquaintance			2.404***
			(4.68)
Friend × No reward			-0.808
			(-1.54)
Acquaintance × No reward			-2.172***
			(-3.49)
Constant	1.917***	1.355***	-0.087
	(8.16)	(6.67)	(-0.35)
Observations	435	435	435

Notes: Logistic regressions with bootstrap standard errors stratified within grade years. Dependent variables are dummies for step count above $3{,}000$ (Columns (1) and (3)) and above $3{,}300$ (Column (2)). Coefficients for anonimity (Columns (1) and (2)) and observer identity (Column (3)) reflect the effect in the *Reward* groups. t-statistics in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

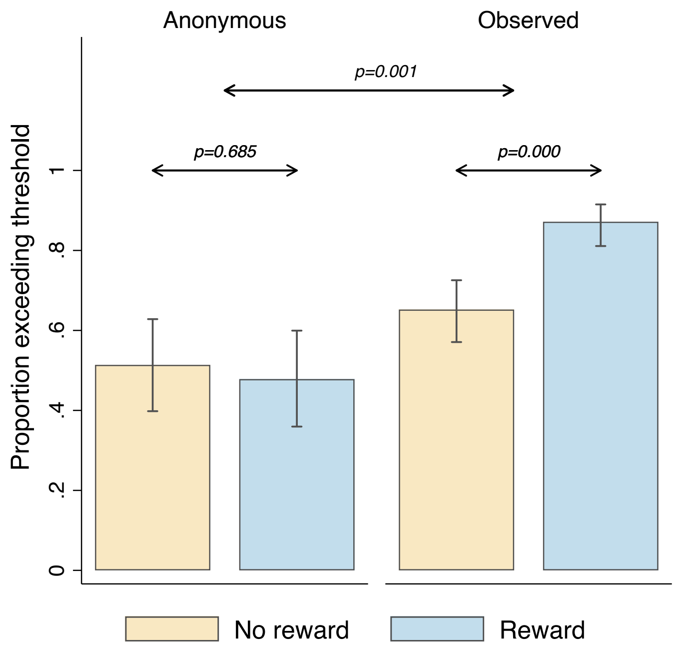

Figure 3: Crowding up.

Notes: Error bars mark the 95% confidence interval based on the regression reported in Column (1) of Table 6.

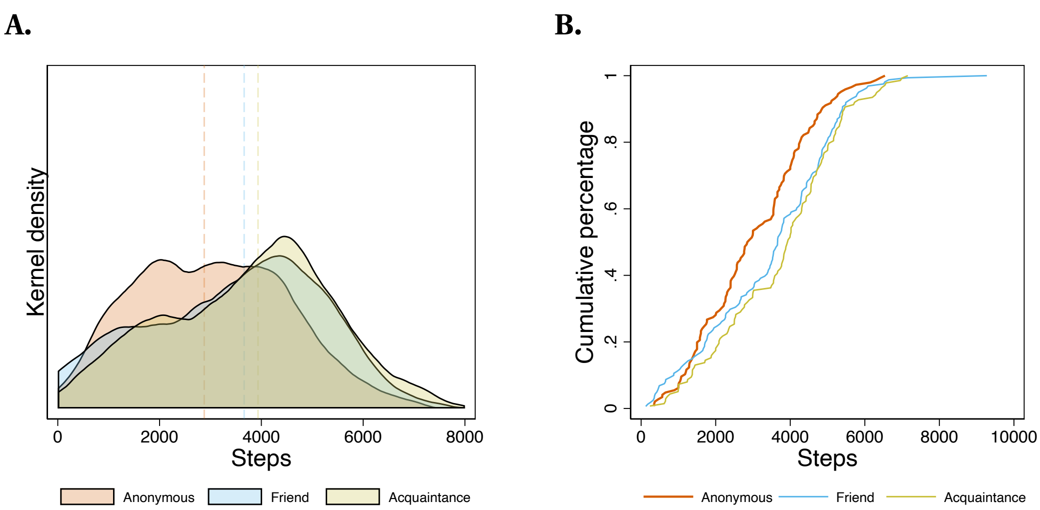

Figure 4: Step-count distributions without reward by observer identity.

Notes: Kernel density function (**A**) and cumulative distribution function (**B**) of the step counts. The vertical dashed lines in Panel **A** mark the medians of the groups.

regression of step count on observer identity interacted with dummies for grades, with bootstrap standard errors stratified by grade years. For the younger students, in the 8th to the 10th grades, observation by an acquaintance has a stronger effect than observation by a friend. While the effect of observation by a friend is never significant, observation by an acquaintance has a significant effect for the 8th grade ($z = 3.74, p < .001$) and for the 10th grade ($z = 2.90, p = .004$). For older students, in the 11th and 12th grades, the picture reverses, with significant effects for observation by a friend but not by an acquaintance.

To see how the effect of observation by acquaintance versus a friend evolves with age, we tested the diff-in-diff effect of observer identity by adjacent grades based on the regression analysis. The differential effect of different observers does not change significantly from the 8th to the 9th grade ($z = -1.10, p = .271$) or from the 9th to the 10th grade ($z = 0.62, p = .538$), but drops significantly (and changes sign) from the 10th to the 11th grade ($z = -2.16, p = .031$). From the 11th to the 12th grade change is again not significant ($z = 0.65, p = .513$). Furthermore, we tested the effects of grade year as a continuous variable and its interactions with observer identity in the regressions reported in Columns (3) and (4) of Table 5. The interaction of grade year with observer identity (friend vs. acquaintance) is highly significant ($z = -2.94, p = .003$).[10] We thus conclude that observer identity has a differential effect, depending on age. For younger adolescents, aged 13 to 15, the familiarity of the observer (μ in our model) plays a larger role, while the older participants care more about the opinions of their close friends (d in the model).[11]

Result 4. *We find support for Hypothesis 3a among younger adolescents, confirming the prediction that observation by an acquaintance may have a stronger effect than observation by a friend.*

We turn next to the effect of observer identity on crowding up. Figure 6 presents the step-count distributions in the *Reward* groups. In the acquaintance group, 71 of the 78 participants (91.0%) walked more than 3,000 steps compared to 65 of 78 (83.3%) in the friend group. The difference is not statistically significant (Fisher's exact $p = .231$). The logistic regression reported in Column (3) of Table 6 yields similar results. The effect of being observed by an acquaintance ($\Delta = 43.2\%, z = 5.97, p < .001$) is stronger than the effect of being observed by a friend ($\Delta = 35.5\%, z = 4.76, p < .001$), but not significantly so ($z = 1.37, p = .172$). The interaction of reward and ob-

[10]This result is robust to excluding the 8th grade from the analysis, with the interaction term only mildly dropping from −382 to 366. The significance level drops, possibly due to the reduced sample size ($z = -2.08, p = .037$).

[11]This can result from either μ decreasing or d increasing with age. More precisely, it is possible that older participants care less about the acquaintance or that the acquaintance is more likely to know the types of older participants, so that the parameters μ and d should be interpreted as the relative weight assigned to a friend compared to the reference weight assigned to an acquaintance.

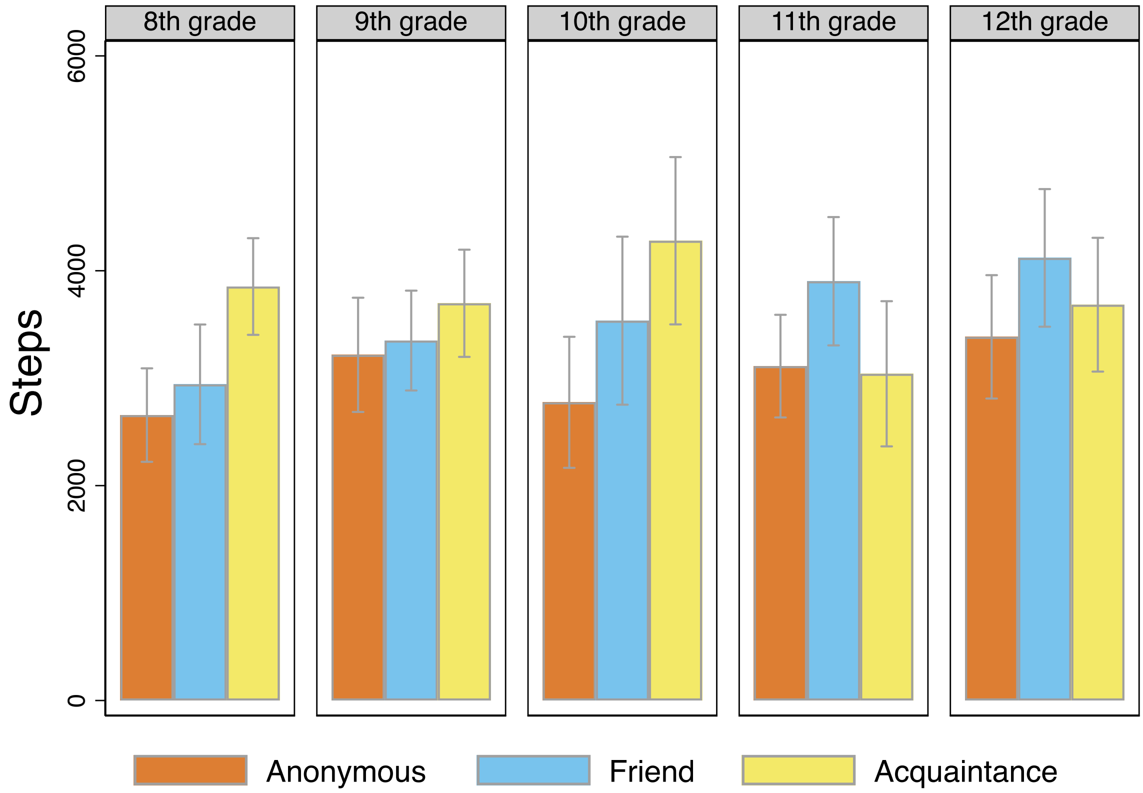

Figure 5: Step count by grade.

Notes: Average step count by grade and observer identity in the No reward treatment. Error bars mark the 95% confidence interval based on an OLS regression of step count on a factorial design including grade, observer identity, and reward with bootstrap standard errors stratified by grade.

Table 7: Step count by grade

	Anonymous vs. Friend	Acquaintance vs. Friend
8th grade	-286	911**
	(-0.80)	(2.42)
9th grade	-133	347
	(-0.37)	(1.00)
10th grade	-760	744
	(-1.54)	(1.34)
11th grade	-790**	-861*
	(-2.04)	(-1.78)
12th grade	-735*	-436
	(-1.69)	(-0.95)
Observations	445	

Notes: Marginal effects of observer identity in the No reward *treatments based on an OLS regression of step count on observer identity interacted with grade year with bootstrap standard errors stratified by grade year. z-statistics in parentheses.* * $p < .10$, ** $p < .05$, *** $p < .01$.

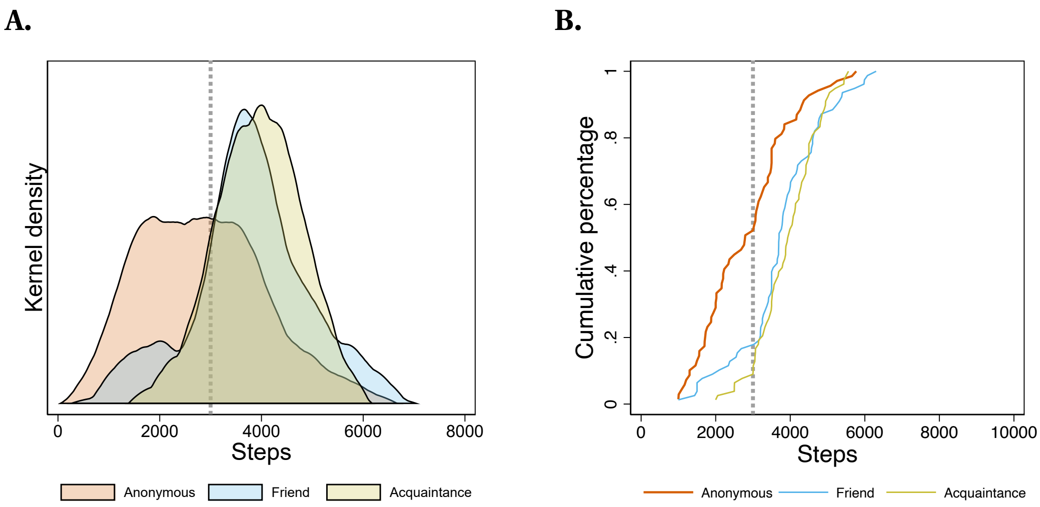

Figure 6: Step-count distributions with reward by observer.

Notes: Kernel density function (**A**) and cumulative distribution function (**B**) of the step counts. The vertical dotted line mark the 3,000-steps threshold.

servation is statistically significant when an acquaintance is observing but not when a friend is. However, this difference can be attributed to the differential effect in the *No reward* treatments.[12]

Recall that we found the observability effect to be stronger for an acquaintance in the 10th grade and for a friend in the higher grades. Accordingly, we should expect to also find a stronger crowding-up effect for an acquaintance among 10th-graders. Figure 7 and Table 8 break down the results by grade based on a logistic regression of a dummy for walking more than 3,000 steps on observer interacted with dummies for grades with bootstrap standard errors stratified by grade years. As expected, participants in the 10th grade who receive a reward are more likely to walk more than 3,000 steps when observed by an acquaintance. The difference between the acquaintance group and the other groups is significant ($z = 3.72, p = .001$ compared to anonymous; $z = 2.34, p = .019$ compared to friend), whereas the difference between the friend and anonymous groups is not ($z = 0.96, p = .338$).

Result 5. *We find support for Hypothesis 3b for 10th-grade participants, with more crowding up when observed by an acquaintance compared to when observed by a friend. There is no apparent difference for older participants, possibly due to a ceiling effect.*

[12]As we saw in Section 4.1, a friend had a stronger effect in the higher grades. Given that participants in these grades walk more than the 10th-graders in the baseline, the effect of observation drives them above the 3,000-steps threshold even without a reward.

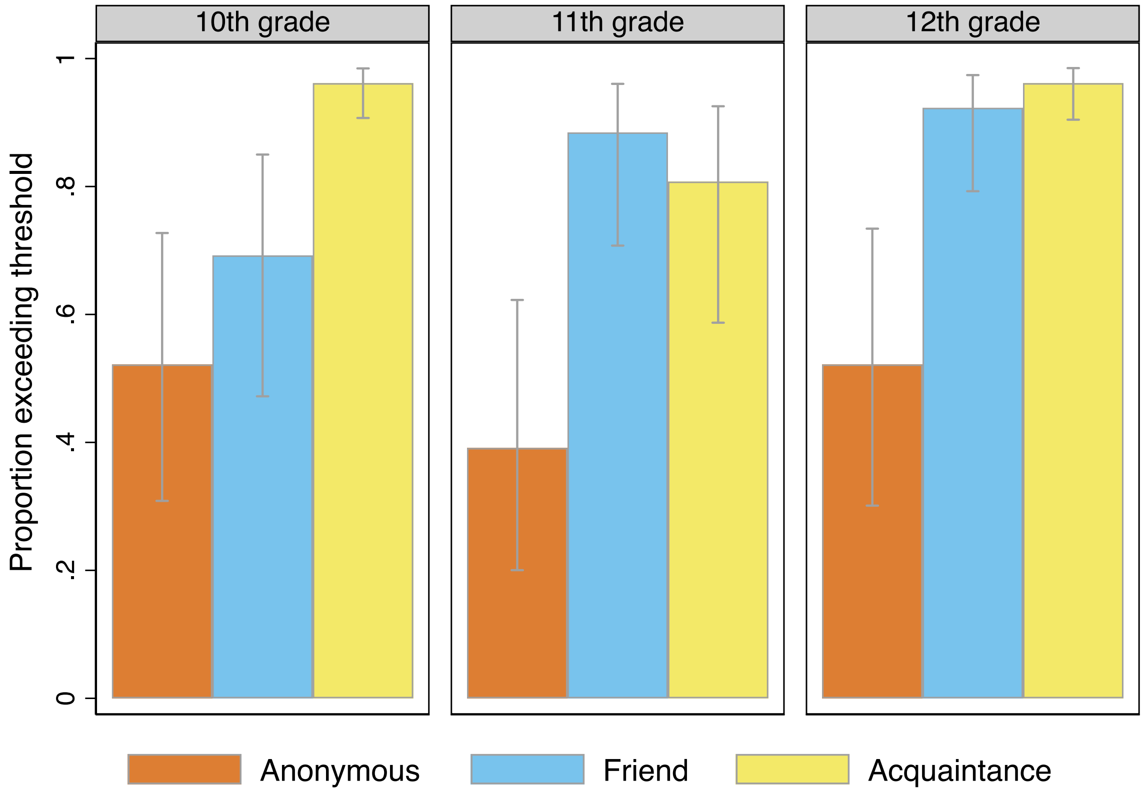

Figure 7: Crowding up by grade.

Notes: Share of participants exceeding 3,000 steps by grade and observer identity in the Reward treatment. Error bars mark the 95% confidence interval based on a logistic regression of exceeding 3,000 steps on grade and observer identity with bootstrap standard errors stratified by grade.

Table 8: Crowding up by grade

	Anonymous vs. Friend	Stranger vs. Friend
10th grade	-0.171	0.269**
	(-0.96)	(2.34)
11th grade	-0.493***	-0.077
	(-4.07)	(-0.63)
12th grade	-0.401***	0.038
	(-2.74)	(0.93)
Observations	225	

Notes: Marginal effects of observer identity in the No reward *treatments based on an OLS regression of step count on observer identity interacted with grade year with bootstrap standard errors stratified by grade year. z-statistics in parentheses.* * $p < .10$, ** $p < .05$, *** $p < .01$.

5. Conclusion

This paper introduces the notion of observer identity into social-image concerns. We build on the notion of social distance to show that observer identity matters in two major dimensions: the importance of the observer's opinions and the observer's prior beliefs. The juxtaposition of the two vis-à-vis social distance imply that social distance interacts with social-image concerns in contrasting ways. Our field experiment provides evidence in both directions. Younger participants respond more to observation by a random peer than to observation by a friend, whereas the older participants appear to care most about the opinions of their close friends. These results illustrate that observer identity matters in non-trivial ways that interact with the characteristics of the observed agent.

While the role of age was not part of our ex-ante hypotheses, there are several indications that the age effects are systematic. The interaction of age and observer identity is consistent across our two hypotheses, tested in separate subsamples. First, younger participants walk more sans reward as well as tend to exceed the reward threshold more when observed by an acquaintance. Second, the stronger effect of an acquaintance diminishes consistently with age. Finally, the results are in line with the psychological literature on peer pressure, which finds that peer pressure is strongest in early adolescents, diminishing across adolescence and flattens at young adulthood (Berndt, 1979; Gardner and Steinberg, 2005; Steinberg and Monahan, 2007). The observation that the older participants care more about the opinions of their close friends mirrors the finding that giving to friends and families among adolescents increases with age whereas giving to strangers does not (Karan et al., 2022).

The paper also contributes to the literature on identifying social-image concerns. We implement the procedure introduced by Birke (2023) and provide novel evidence for crowding up in the field. Crowding up reflects social-image concerns while avoiding potential confounds such as social facilitation (Zajonc, 1965). In our setting, it serves the additional purpose of pinning down the signaled dimension. In general, people may wish to signal ability as well as preferences. In the experiment, participants may exert more effort when being observed to signal their physical ability. The crowding-up effect, however, relies on the reward undermining the signal value of below-threshold performance, which is relevant for signaling intentions but not abilities.

Understanding the role of observer identity for designing interventions that leverage social-image concerns to promote normative behaviors. Contemporary technology allows designers to target observers with directed information. This paper suggests that practitioners should consider the audience and its relation to the target population. While here we focus on social distance and the moderating effect of age, the message is more broad, as different dimensions may be relevant in different contexts and with different populations.

References

Steiger, Eva-Maria and Ro'i Zultan (2014). See no evil: information chains and reciprocity. *Journal of Public Economics* 109, pp. 1–12.

Adriani, Fabrizio and Silvia Sonderegger (2019). A theory of esteem based peer pressure. *Games and Economic Behavior* 115, pp. 314–335.

Alpizar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman (2008). Anonymity, reciprocity, and conformity: evidence from voluntary contributions to a national park in costa rica. *Journal of Public Economics* 92.(5-6), pp. 1047–1060.

Andreoni, James and Ragan Petrie (2004). Public goods experiments without confidentiality: a glimpse into fund-raising. *Journal of public Economics* 88.(7-8), pp. 1605–1623.

Ariely, Dan, Anat Bracha, and Stephan Meier (2009). Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American economic review* 99.(1), pp. 544–55.

Asch, Solomon E. (2016). Effects of group pressure upon the modification and distortion of judgments. In: *Groups,leadership and men.* Ed. by H. Guetzkow. Carnegie Press, pp. 177–190.

Ashraf, Nava, Oriana Bandiera, and Scott S. Lee (2014). Awards unbundled: evidence from a natural field experiment. *Journal of Economic Behavior & Organization* 100, pp. 44–63.

Asulin, Yamit, Yuval Heller, Nira Munichor, and Ro'i Zultan (Dec. 2023). Social image, observer identity, and crowding up: supplementary materials. URL: osf.io/qxbya.

Bašić, Zvonimir and Simone Quercia (2022). The influence of self and social image concerns on lying. *Games and Economic Behavior* 133, pp. 162–169.

Bénabou, Roland and Jean Tirole (2006). Incentives and prosocial behavior. *American economic review* 96.(5), pp. 1652–1678.

Berndt, Thomas J. (1979). Developmental changes in conformity to peers and parents. *Developmental psychology* 15.(6), pp. 608–616.

Bicchieri, Cristina, Eugen Dimant, Simon Gächter, and Daniele Nosenzo (2022). Social proximity and the erosion of norm compliance. *Games and Economic Behavior* 132, pp. 59–72.

Birke, David J. (2023). Anti-bunching: a new test for signaling motives in prosocial behavior.

Bond, Charles F. and Linda J. Titus (1983). Social facilitation: a meta-analysis of 241 studies. *Psychological bulletin* 94.(2), p. 265.

Bursztyn, Leonardo, Georgy Egorov, and Robert Jensen (2019). Cool to be smart or smart to be cool? understanding peer pressure in education. *The Review of Economic Studies* 86.(4), pp. 1487–1526.

Bursztyn, Leonardo and Robert Jensen (2015). How does peer pressure affect educational investments? *The quarterly journal of economics* 130.(3), pp. 1329–1367.

— (2017). Social image and economic behavior in the field: identifying, understanding, and shaping social pressure. *Annual Review of Economics* 9, pp. 131–153.

Cameron, Andrew M., Allan B. Massie, Charles E. Alexander, B. Stewart, Robert A. Montgomery, N. R. Benavides, GD Fleming, and Dorry L. Segev (2013). Social media and organ donor registration: the facebook effect. *American Journal of Transplantation* 13.(8), pp. 2059–2065.

Clendenen, Vanessa I., C. Peter Herman, and Janet Polivy (1994). Social facilitation of eating among friends and strangers. *Appetite* 23.(1), pp. 1–13.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier (2012). Testing for altruism and social pressure in charitable giving. *The quarterly journal of economics* 127.(1), pp. 1–56.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao (2016). Voting to tell others. *The Review of Economic Studies* 84.(1), pp. 143–181.

Exley, Christine (2018). Incentives for prosocial behavior: the role of reputations. *Management Science* 64.(5), pp. 2460–2471.

Friedrichsen, Jana and Dirk Engelmann (2018). Who cares about social image? *European Economic Review* 110, pp. 61–77.

Funk, Patricia (2010). Social incentives and voter turnout: evidence from the swiss mail ballot system. *Journal of the European economic association* 8.(5), pp. 1077–1103.

Gardner, Margo and Laurence Steinberg (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental psychology* 41.(4), pp. 625–635.

Graf, Caroline, Bianca Suanet, Pamala Wiepking, and Eva-Maria Merz (2023). Social norms offer explanation for inconsistent effects of incentives on prosocial behavior. *Journal of Economic Behavior & Organization* 211, pp. 429–441.

Karan, Maira, Lee Lazar, Carrianne J. Leschak, Adriana Galván, Naomi I. Eisenberger, Jessica P. Uy, Macrina C. Dieffenbach, Eveline A. Crone, Eva H. Telzer, and Andrew J. Fuligni (2022). Giving to others and neural processing during adolescence. *Developmental cognitive neuroscience* 56, p. 101128.

Karing, Anne (2023). Social signaling and childhood immunization: a field experiment in Sierra Leone. *mimeo.*

Kraft-Todd, Gordon, Erez Yoeli, Syon Bhanot, and David Rand (2015). Promoting cooperation in the field. *Current Opinion in Behavioral Sciences* 3, pp. 96–101.

Li, Mei, Jin Li, Guanfei Zhang, Wei Fan, Yiping Zhong, et al. (2022). Interpersonal distance modulates the influence of social observation on prosocial behaviour:

an event-related potential (ERP) study. *International Journal of Psychophysiology* 176, pp. 108–116.
McKelvey, Wendy and Nancy H. Kerr (1988). Differences in conformity among friends and strangers. *Psychological Reports* 62.(3), pp. 759–762.
Meyer, Christian Johannes and Egon Tripodi (2021). Image concerns in pledges to give blood: evidence from a field experiment. *Journal of Economic Psychology* 87, p. 102434.
Ministry of Education (2014). Accessed: November 25, 2023. Ministry of Education. URL: https://www.gov.il/he/departments/news/personal-commitment.
Perez-Truglia, Ricardo and Guillermo Cruces (2017). Partisan interactions: evidence from a field experiment in the united states. *Journal of Political Economy* 125.(4), pp. 1208–1243.
Rege, Mari and Kjetil Telle (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of public Economics* 88.(7-8), pp. 1625–1644.
Rogers, Todd, John Ternovski, and Erez Yoeli (2016). Potential follow-up increases private contributions to public goods. *Proceedings of the national academy of sciences* 113.(19), pp. 5218–5220.
Steinberg, Laurence and Kathryn C Monahan (2007). Age differences in resistance to peer influence. *Developmental psychology* 43.(6), p. 1531.
Tice, Dianne M., Jennifer L/ Butler, Mark B. Muraven, and Arlene M. Stillwell (1995). When modesty prevails: differential favorability of self-presentation to friends and strangers. *Journal of personality and social psychology* 69.(6), pp. 1120–1138.
Wagner, Hugh L. and Jayne Smith (1991). Facial expression in the presence of friends and strangers. *Journal of Nonverbal Behavior* 15, pp. 201–214.
Yoeli, Erez, Moshe Hoffman, David G Rand, and Martin A Nowak (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences* 110.(supplement_2), pp. 10424–10429.
Zajonc, Robert B. (1965). Social facilitation: a solution is suggested for an old unresolved social psychological problem. *Science* 149.(3681), pp. 269–274.

A. Proof of Proposition 1 and Corollaries 1–3

It is immediate from the payoff structure that in all cases type θ_2 plays action a_2, type θ_1 plays either a_1 or a_2, and type θ_0 plays either a_0 or a_1 (type θ_0 never plays action a_2 due to the assumption that $u_{02} < u_{01} - 2$). Let p_1 be the equilibrium probability of an agent of type θ_0 playing action a_1, and let p_2 be the equilibrium probability of an agent of type θ_1 playing action a_2.[13]

Acquaintance and No reward We begin by analyzing the setup with no reward and observability by an acquaintance. The assumption that $c_1 < 1$ rules out the case that an agent of type θ_0 always plays action a_1. The assumption that $\frac{q_1}{q_0+q_1} \leq \frac{1}{2} < c_1$ rules out the case that an agent of type θ_0 always plays a_1. Thus, agent of type θ_0 must mix between actions a_0 and a_1. The indifference between these two options implies that:

$$c_1 = E\left(\theta|a_1\right) = \frac{(1-p_2)\, q_1}{(1-p_2)\, q_1 + p_1 q_0} \Leftrightarrow p_1 = \frac{(1-p_2)\, q_1\, (1-c_1)}{c_1 q_0}. \tag{1}$$

The assumption that $c_2 < d\,(1-\mu) < 1$ rules out the case that an agent of type θ_1 always plays a_1. The assumption that $\frac{q_2}{q_1+q_2} < c_2$ rules out the case that an agent of type θ_1 always plays a_2. Thus, agent of type θ_1 (who cares about reputation) must mix between actions a_1 and a_2. The indifference between these two options implies that:[14]

$$c_2 = E\left(\theta|a_2\right) - E\left(\theta|a_1\right) = 1 + \frac{q_2}{q_2 + p_2 q_1} - c_1 \Leftrightarrow p_2 = \frac{q_2\,(2-c_1-c_2)}{q_1\,(c_2+c_1-1)}. \tag{2}$$

Observe that p_2 is decreasing in c_2. Substituting p_2 in Equation (1) and simplifying yields:

$$p_1 = \frac{1-c_1}{c_1 q_0}\left(q_1 - \frac{(2-c_1-c_2)\, q_2}{(c_1+c_2-1)}\right). \tag{3}$$

[13] More formally, ϵ of the agents of type θ_0 who do not care about reputation, always plays a_0. The remaining $1-\epsilon$ agents of type θ_0 play a_1 with probability $\frac{p_1}{1-\epsilon}$. This induces a total probability of p_1 of an agent of type θ_0 playing a_1. Similarly, ϵ of the agents of type θ_1 always play a_1. The remaining $1-\epsilon$ agents of type θ_1 play a_2 with probability $\frac{p_2}{1-\epsilon}$. This induces a total probability of p_2 of an agent of type θ_1 playing a_2.

[14] More formally, ϵ of the agents of type θ_1 who do not care about reputation, always plays a_1. The remaining $1-\epsilon$ agents of type θ_0 play a_1 with probability $\frac{p_1}{1-\epsilon}$. This induces a total probability of p_1 of an agent of type θ_0 playing a_1.

The equilibrium expected level of effort is given by:

$$\begin{aligned} E_\sigma(a_i) &= p_1 q_0 + q_1(1+p_2) + 2q_2 \\ &= \frac{(1-p_2)\, q_1 (1-c_1)}{c_1} + q_1(1+p_2) + 2q_2 \\ &= q_1 + 2q_2 + \frac{(1-p_2)\, q_1 (1-c_1)}{c_1} + q_1 p_2 \\ &= q_1 + 2q_2 + q_1\left((1-p_2)\frac{1-c_1}{c_1} + p_2\right), \end{aligned}$$

which is decreasing in c_1 and increasing in p_2 due to the assumption that $c_1 \in (0.5, 1)$.

Friend and No reward The equilibrium in the setup without reward and with observability by a friend is characterized by the same indifference equations as in (1) of p_1 and p_2, except that $E(\theta|a_1)$ and $E(\theta|a_2)$ are replaced by $v \cdot E(\theta|a_1)$ and $v \cdot E(\theta|a_2)$, where $v = d \cdot (1-\mu)$. This is equivalent to replacing c_1 and c_2 with $\frac{c_1}{v}$ and $\frac{c_2}{v}$, respectively, and obtaining the following equations for p_1 and p_2:

$$\begin{aligned} p_2 &= \frac{q_2(2v - c_1 - c_2)}{q_1(c_2 + c_1 - v)}, \\ p_1 &= \frac{(1-p_2)\, q_1 (v - c_1)}{c_1 q_0} = \frac{v - c_1}{c_1 q_0}\left(q_1 - \frac{(2v - c_1 - c_2)\, q_2}{(c_1 + c_2 - v)}\right). \end{aligned}$$

Observe that $E_\sigma(a_i)$ is decreasing in c_1 and increasing in p_2 in this setup as well due to the assumption that $c_1 < v \Rightarrow \frac{c_1}{v} \in (0.5, 1)$. This implies that changing observability from an acquaintance to a friend (which increases c_1 and c_2), decreases p_2 and $E_\sigma(a_i)$, which implies part (1) of Corollary 3.

Large reward ($b \geq c_1$) Next, we analyze the setups with a reward. Assume first, that $b > c_1$. This implies that an agent of type θ_0 always plays a_1. This, in turn, implies that an agent of type θ_1 (when being observed by either a friend or an acquaintance) always plays a_2 because the cost of playing action a_2 is smaller than the reputation payoff of playing a_2:

$$c_2 < v < v\left(1 + \frac{q_2}{q_2 + q_1}\right) < 1 + \frac{q_2}{q_2 + q_1}.$$

Thus, with a large reward and observability by either a friend or an aquiatnace, type θ_0 always plays a_1, while almost all of type θ_1 agents (the share of $1-\epsilon$ of them who care about reputation) and all type θ_2 agents play a_2. Thus, the large reward case satisfies Corollaries 2–3.

Small reward ($b < c_1$) Next, assume that $b < c_1$. This implies that an agent of type θ_0 (who cares about reputation) is indifferent in the unique equilibrium between playing a_1 and a_2. The indifference equation for an agent of type θ_0 is the same as in (1) except that c_1 is replaced by $c_1 - b$ for an observing acquaintance:

$$c_1 - b = E\left(\theta | a_1\right) = \frac{\left(1 - p_2\right) q_1}{\left(1 - p_2\right) q_1 + p_1 q_0} \Leftrightarrow p_1 = \frac{\left(1 - p_2\right) q_1 \left(1 - c_1 + b\right)}{\left(c_1 - b\right) q_0}, \qquad (4)$$

and by $\frac{c_1 - b}{v}$, or for an observing friend:

$$c_1 - b = v \cdot E\left(\theta | a_1\right) = \frac{v \cdot \left(1 - p_2\right) q_1}{\left(1 - p_2\right) q_1 + p_1 q_0} \Leftrightarrow p_1 = \frac{\left(1 - p_2\right) q_1 \left(1 - \frac{c_1 - b}{v}\right)}{\frac{c_1 - b}{v} q_0}$$

Agent of type θ_1 (who cares about reputation) strictly prefers playing action a_2 (which implies that $p_2 = 1 - \epsilon$) if

$$c_2 < 1 + \frac{q_2}{q_2 + q_1} - \left(c_1 - b\right),$$

for observability by an acquaintance and

$$c_2 < v\left(1 + \frac{q_2}{q_2 + q_1} - \left(c_1 - b\right)\right),$$

for observability by a friend.

If the opposite inequality holds, then type θ_1 plays a_2 with probability p_2, which is determined by the following indifference equation:

$$c_2 = E\left(\theta | a_2\right) - E\left(\theta | a_1\right) = 1 + \frac{q_2}{q_2 + p_2 q_1} - \left(c_1 - b\right) \Leftrightarrow p_2 = \frac{q_2\left(2 + b - c_1 - c_2\right)}{q_1\left(c_2 + c_1 - 1 - b\right)},$$

for observability by an acquaintance, and

$$p_2 = \frac{q_2\left(2v + b - c_1 - c_2\right)}{q_1\left(c_2 + c_1 - v - b\right)}$$

for observability by a friend. Observe that in all cases either p_2 is larger in the setup of observability with an acquaintance or $p_2 = p_1 = 1 - \epsilon$, which implies part (2) of Corollary 3.