HOW TO CHOOSE YOUR VICTIM*

KLAUS ABBINK, GÖNÜL DOĞAN[1]

This Version: 6 March 2014

We introduce the experimental mobbing game. Each player in a group has the option to nominate one of the other players or nominate no one. If the same person is nominated by all other players, he loses his payoff and the bullies gain. We conduct three sets of experiments to study the effects of monetary gains, focality and fear of being bullied. We find that subjects frequently coordinate on a victim, even for only modest gains. Higher gains make mobbing more likely. Richer and poorer players are equally focal. We find no evidence that fear of becoming the victim explains bullying.

Jel Codes: C92, D03, D70

* Gönül Doğan gratefully acknowledges the financial help of the Research Priority Area Behavioral Economics of the University of Amsterdam.

1 Abbink: Department of Economics, Monash University, Wellington Road, Clayton, VIC 3800, Australia. klaus.abbink@monash.edu. Corresponding Author: Doğan: CREED, University of Amsterdam, Roetersstraat 11, 1018WB Amsterdam, the Netherlands. gonul2517@gmail.com.

1. Introduction

The commonly accepted definition of bullying in the psychology literature involves three components: intentional behavior to cause harm, repeated interactions, and an imbalance of power between the perpetrator(s) and the victim (see Olweus, 1993, 1994; Gladden et al., 2014). Bullying, when measured via questionnaires by focusing on these three aspects, is found to be highly prevalent both at schools, in the workplace, and recently also in social networks on the internet. The estimates of students involved as a victim, perpetrator, or both range from 20% to 56% in the US[2], whereas between 3% and 15% of all employees are estimated to be victims of serious bullying in the workplace (Zapf et al. (2010)). Moreover, bullying inflicts severe consequences such as depression, loneliness, anxiety, and low self-esteem on the victims. Due to loss of productivity or absence from work, the direct costs of bullying in the workplace are enormous. As an example, in 2010, the Australian Productivity Commission estimated the costs to employers to be between 6 and 36 billion Australian dollars per year; the latter figure corresponds to roughly 2.5% of the country's GDP. Finding strategies against bullying is therefore important, and understanding its causes is a primary prerequisite for it.

In the field, bullying behaviour is notoriously hard to observe, which explains the wide margins in the estimates cited above. Much of it takes place below the surface, not least because victims often internalize it, and the surrounding culture lets perpetrators get away with their behaviour. Questionnaire studies deliver valuable insights over the incidence rate as well as the characteristics of victims and to some extent the bullies, however, these studies do not allow full control over the accuracy of the data, and therefore are rather silent over the causes of bullying, and the relative importance of different dimensions of bullying. To complement

[2] See Centers for Disease Control and Prevention, (2011), Borowsky et al. (2013), or Kowalski and Limber (2013).

this existing literature, we suggest a novel approach to the analysis of bullying. In a laboratory experiment, we can, under controlled conditions, construct a game that captures essential features of a potential bullying scenario. By varying the conditions across treatments, we can identify factors that contribute to bullying or prevent it to occur, and test hypotheses about the motivations underlying the observed behaviour.

In this study we look at mobbing, a particular form of bullying in which an individual is harassed by a group of tormentors. This is a more sophisticated setting than a one-on-one encounter, as it requires the coordination of several individuals, some of whom may be reluctant to take part. In fact, some might only participate to avoid challenging the group and not become a victim themselves. Situations can easily arise in which groups become locked in – bullying continues because no one can simply leave the group without risk. We capture this group dynamic in a simple experimental game, which we call the *mobbing game*. In a group of players, each player can, but is not forced to, nominate a victim among the group members. If players nominate different individuals or none, then there is no victim, and players receive their default equal payoffs. If all other players nominate the same individual, then this individual becomes the victim, his payoff is taken away and the bullies receive an additional payoff. This game is repeated with the same group of players over a fixed number of periods. What we obtain is a new, darker incarnation of an experimental coordination game – players do not coordinate to find the best common choice, but to conspire against a fellow group member. As such, our game captures the three crucial aspects of bullying employed in the psychology literature: the intentions are clear and hurtful to the victim, actions are repeated, and the victim is powerless against the perpetrators. Using this new paradigm, we can study the coordination problem involved in mobbing, and thereby address some important questions: (1) How do individuals coordinate on choosing a victim, (2) Who is most likely to become a victim, and (3) What motivates individuals to choose a victim?

We conduct three sets of experiments, comprising altogether eight treatments. The first set uses a symmetric setting without any focal players or coordination devices. In this first set, we want to investigate the 'price' of being part of the mob. Therefore, we keep the victim's payoff loss constant, and we vary the benefits to the bullies from mobbing. We create three situations that let us investigate the prevalence of bullying: when bullying is detrimental to efficiency, when it is redistributive, and when it is efficiency enhancing. Although the first two situations are more representative of bullying at schools, there are many occasions in the workplace in which coordinating on a victim actually enhances efficiency. A case in point is a group of colleagues eliminating another colleague who is a competitor for a promotion or a project: The bullies' total payoff increase is more than the cost to the victim due to increased chances of promotion and less competition for the bullies[3].

In the first set of experiments, we observe that, players find it remarkably easy to pick a victim, and are surprisingly inclined to do it even if the gains for the bullies are modest. To be more precise, low payoffs from bullying leads to little bullying, and high payoffs from bullying leads all groups to engage in bullying. When the payoffs from bullying are modest, i.e., when bullying is just a redistribution of payoffs, we observe approximately half of the groups engaging in it. Therefore, in the rest of the experiments, we focus on varying some aspects of the redistributive bullying case.

In the second set of experiments, we investigate the effect of focality, i.e., being different. It is a widely accepted finding in the psychological literature on bullying that victims are almost always different in some aspect. To investigate the importance of being different in the mobbing game, and building up on the results of the first set of experiments that show the ease of coordination, we introduce asymmetry by making one of the players either richer or poorer than the others. To achieve comparability both within the second set of experiments as well

[3] Efficiency gain in the group does not mean that there is an efficiency gain at the organization level.

as with our baseline (modest payoff gain with redistribution of payoffs) game, we keep the victim's cost constant, and we also keep the additional payoff from bullying the same for any player. Thus, in the second set of experiments, bullying is again redistributive. In addition to investigating the effect of focality, the second set of experiments further let us study whether commonly used social preference theories have any predictive power.

In the second set of experiments, we find that the player who is different is by far the most likely to become the victim, no matter whether the odd one out is richer or poorer by default. Moreover, and contrary to commonly used social preferences model predictions, the existence of a poorer player does not decrease the overall bullying rate, whereas the existence of a richer player increases it.

Finally, we designed the third set of experiments to disentangle possible motives behind bullying. Our experiments are stripped of a social context, therefore, and unlike real-life bullying, the motives for being part of the mob are mostly limited to greed and fear[4]. To study the effect of fear, we introduce a safe player to the baseline game, that is, a player who can be chosen but without any payoff consequences. A safe player has, thus, nothing to be afraid of. Moreover, there cannot be a mob without the safe player being in it. If fear of being a victim is the driver of bullying behaviour, we would expect that safe players behave differently than non-safe players, and that their introduction decreases mobbing rates. The results of this third set of experiments show that the pursuit of gain drives bullying, not the fear of becoming the victim oneself.

To our knowledge, there are no studies that propose a similar game to ours. The closest studies to the present work are coordination games, which are due to the work of Van Huyck et al. (1990, 1991). In these games, players have to coordinate on one of multiple Pareto ranked

[4] It is conceivable that observing that there are others in one's group who behave in a certain way also has an effect on the individual behavior. We will say something about this effect in section 5.

equilibria. In their experiments subjects did poorly – they typically coordinated on the worst equilibrium. Such coordination failure has been replicated in various experiments and is found to be dependent on the attractiveness of the efficient outcome, the time horizon, the group size, the deviation costs, refined action space, and communication (for a meta-analysis on coordination experiments see Devetag and Ortmann (2007)). Cooper et al. (1990, 1992) introduced the stag-hunt game, in which one equilibrium is Pareto dominant and the other is risk-dominant, and find frequent coordination failure. Isoni et al. (2013) and Poulsen et al. (2013) analyse solutions to coordination problems with respect to their saliency. In all these studies, there is no victim; all players benefit from successful coordination. The model most akin to ours is the collective resistance game by Cason and Mui (2007), in which two followers must coordinate if they want to challenge a transgressing leader. The results show that followers are very successful in coordinating their resistance, unless the leader is allowed to use divide-and-conquer strategies, i.e. target individual followers rather than both. The endogenous victim selection, which is at the heart of our study, is not part of their research agenda.

Lastly, our study contributes to the newly emerging literature on the effect of groups on decision making. Studies by Falk and Szech (2013a, 2013b) have shown that people behave differently when confronted with the same decision individually or in a group. Through a series of experiments that involves trading the life of a mouse, they showed that being in a group dilutes responsibility, and lets people behave less morally. Our study diverges from their set-up in two crucial ways: in our game all players are pivotal in decision-making, and in all our treatments the price of being part of a mob is fixed which limits social learning. Finally, the principal question that we are trying to answer in this paper is not establishing the difference between group behaviour and individual behaviour, but rather the conditions under which a certain type of negative group behaviour emerges.

2. The experimental design

All experiments were conducted at the Center for Research in Experimental Economics and political Decision making (CREED) of the University of Amsterdam. Subjects were recruited via the online recruitment system of CREED and were mostly undergraduates from a wide variety of majors. Each subject could participate in only one session, and all treatments were across subjects, i.e. in each session only one treatment was run. There were 8 treatments conducted across 21 sessions. Average pay was 13.24 Euros including 7 Euros show-up fee. The experiments lasted between 40 minutes and 1 hour 15 minutes including the time spent on reading the instructions.

The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007). The experiments were in English, the language of instruction at the university. The instructions used neutral language; a copy of the instructions for the baseline treatment as explained below is included in the Supplementary Data. It was explained in the instructions that each group consisted of four players (or five, in treatments we describe later) who were randomly selected at the beginning of the experiment and that the group composition stayed the same throughout the whole experiment (partners design). Subjects had no way of knowing which of the other participants were in their group. In the instructions and the experiment, the players were denoted as M, T, G, and P, a seemingly random set of consonants.[5] Player labels were randomly assigned at the beginning of the experiment and remained fixed throughout the experiment.[6] Subjects were informed that the task would be repeated for 20 periods and that their final earnings comprised of the total points they earned in the experiment converted

[5] The labels are a hidden homage to the inmates Mather, Travers, Greenhill and Pearce, who escaped from a Tasmanian prison camp in a group of eight in 1822, only to get lost in the forest. When food ran out, the four conspired to apply the Custom of the Sea to the others. When no one else was left, they turned to killing and eating one another, until only Pearce survived. All victims were chosen in decidedly non-random ways. This story is one of the great Australian foundation myths, and it was an inspiration for this study (for a dramatic reconstruction see *Van Diemen's Land* (2009)). We are confident that none of our Amsterdam subjects made that connection.

[6] In the treatments with five players we also used the label R, with no historical reference.

at a rate of 1.25 euro cents per point plus the show up fee. The earnings scheme depicts the idea of repeated and accumulated damage done to another party, and is a representation of the real life mobbing situation we are interested in studying. After the instructions were finished, subjects played one practice period with full feedback in order to familiarise them with the procedure and the screens.

Since subjects received no feedback from other groups and the groups remained the same, each group constitutes a statistically independent observation.

3. The first set of experiments: The symmetric baseline

The basic game is very simple: Each member of a group of four players[7] decides simultaneously to nominate one of the other players in their own group, or not to nominate anyone. Nominating as such is neither costly nor beneficial. If three players nominate the same fourth player, then the nominee loses all of his earnings whereas the nominators earn additional money. The game is repeated for 20 periods with the same players. After each period, subjects are given feedback about the number of nominations each player received, but not from whom.

In the three treatments of the first set of experiments, all four players are identical. They have the same payoffs in absence of bullying[8] (which we refer to as the default payoff), and the same prospective payoffs in case that they become victims or bullies. Table I shows the payoff distributions. In all three treatments, all players receive 24 tokens if there is no bullying, i.e. if three players do not nominate the fourth player either by refraining from nominating a victim or by miscoordination. Further, in all three treatments the victim receives zero payoff. The treatments differ in the payoffs the bullies receive in case of successful coordination, and

[7] Zapf et al. (2010) find that the typical number of bullies in mobbing cases is between two and four – our choice of groups of four , hence three bullies and one victim, is well in line with this figure.

[8] In the following, we will use mobbing and bullying interchangeably. Mobbing is generally used as a subcategory – a group harasses an individual, while bullying includes other forms, like harassment by an individual.

hence in the efficiency of bullying (measured, as usual in experimental economics, as the sum of the payoffs of all players in a group). In the High treatment, bullying is efficiency-enhancing, i.e. the sum of all players' payoffs is greater than the sum of all default payoffs. In the Medium treatment bullying is efficiency-neutral, while in the Low treatment the victim's loss far outweighs the minor gains successful bullies can achieve.

Table I here

Assuming own-payoff maximisation, the game theoretic analysis of the one-shot mobbing game is straightforward. Any outcome in which no player receives two nominations (more generally n-2 nominations if n is the number of players) is an equilibrium outcome. In all these constellations no player can unilaterally improve their payoff. Only the case of one player being nominated exactly by two other players is not an equilibrium outcome, since the remaining non-victim can join the bullies and improve his payoff. Hence coordination on a victim is an equilibrium just as well as non-coordination.

Figure I here

The left-hand panel of Figure I shows the relative frequency of successful bullying, i.e. the fraction of groups which coordinated on the same victim, over the 20 periods of the experiment. The right-hand panel shows the frequency of nominations, i.e. the fraction of individuals who nominated a player, regardless of whether coordination could be achieved.[9] Table II

[9] Note that nominating a victim is inconsequential if there are less than two other nominations on the same victim, so we cannot say for certain that every nomination is carried out with the intention of forming a mobbing

below shows the overall average rates of successful bullying across all groups and periods. Note that the overall rate of successful bullying is an average over the frequency of successful bullying over all the periods; the bullying rates are higher if we exclude the first periods that are needed for coordination. We can make three main observations. First, bullying becomes the more frequent the higher the individual gains from it are. The overall frequencies of successful mobbing are 86.0% in High, 47.3% in Medium, and 13.5% in Low. We can reject the null hypotheses of no difference for all pairwise comparisons.

Table II here

A second observation is that if there is a will to choose a victim, coordination is remarkably easy to accomplish even though all players are identical and have no coordination devices at their disposal. Coordination happens quickly: If we look at the groups in which there existed a victim for at least three periods[10], we see that it took on average 2.8, 3.9 and 5.0 periods in the High, Medium and Low treatments, respectively. Such quick coordination on a victim may seem surprising, but it turns out that in the repeated game context, the coordination problem is not as hard as it seems. If every player arbitrarily nominates another player, it is actually quite likely that one player receives two nominations by chance. It then becomes easy for the remaining non-victim to jump on the bandwagon. Once full coordination is achieved, the incentive for any bully to jump off again is limited – it not only leads to an immediate loss of

group. Thus, while nominating behaviour is certainly informative, we focus predominantly on successful bullying in our analysis.

[10] There are 10, 9, and 3 such groups in the High, Medium and Low treatments, respectively. Our choice of a threshold of three periods is of course somewhat arbitrary.

earnings, but also bears the risk of becoming the next victim. The victim himself is defenceless.[11]

Finally, we can observe that subjects show a surprisingly strong propensity to engage in bullying. Even in the Low treatment, where the gain from bullying is minuscule compared with the damage it does to the victim, on average about half of all subjects nominated a victim. This indifference towards a fellow group member's distress is somewhat at odds with traditional experimental findings that subjects care much about fairness and equality (see the vast literature on dictator games since Hoffman et al (1994) and Forsythe et al. (1994)). Two motives could drive the high frequency of bullying. Subjects may be motivated by the additional money they can gain by coordinating on a victim. Or alternatively, they might nominate a player in order to avoid becoming a victim themselves. Note that it requires all three non-victims to establish mobbing, thus each player has veto power against mobbing any other player. If a player does not nominate, and the other players are inclined to bully, then this player might fear that he eventually becomes the only possible victim. Hence, even in absence of material motives, fear of getting into this situation may drive subjects towards nominating.[12] In our third set of experiment (section 5) we design treatments that disentangle these possible motives.

In the High treatment, it could be a sensible group behaviour to rotate the victim's role. This would capture the efficiency gain without victimising a particular individual. However, we hardly observe any such behaviour. All groups in this treatment coordinated on one victim more than half of the periods, and 8 out of 10 groups coordinated on the same victim at least

[11] It is of course possible that the easy coordination was due to the labelling of players, which may have made one player particularly focal. However, our data do not show a strong bias. On average across all three treatments, player M received 0.77 nominations, player T 0.58, player G 0.74 and player P 0.98 (we look at nominations because focality through labelling should express itself in more nominations). These differences are too small to explain the quick coordination we observe.

[12] Abbink and de Haan (2013) find fear to be a powerful driver of behaviours that are hurtful to others.

80 percent of the time. In the other treatments, no collective gain can be achieved through taking turns, and as expected, we do not see much of it either. In Medium, 5 out of 13 groups coordinated on the same victim 80 percent of the time; in Low, there were two groups that coordinated on the same victim in 7 and 13 periods, respectively. All these figures correspond well to the overall mobbing rates, so if there is mobbing, the victim's role typically falls to the same person.

Finally, we would like to put a note on the predictive power of the social preferences models. Take for example one of the most commonly used social preferences model, the model of Fehr and Schmidt (1999)[13]. In their model people dislike earning more than others, and also others earning more than them. Nonetheless, people are assumed to prefer earning one unit more to earning one unit less even if this means increasing the earnings gap with another person. This assumption has been confirmed by many experiments testing this model, such as the study of Yang, Onderstal and Schram (2013). Therefore, such a model's prediction would be to observe full bullying in any treatment in which bullying results in the redistribution of victim's payoffs or gives higher payoffs than that[14].

All our treatments except the Low treatment, including the second and third set of experiments satisfy this requirement. The rationale is as follows. Any mobbing game with redistribution of payoffs such as in the Medium treatment requires a unit of gain for oneself to be compared to the loss of three units for the victim *weighted* by the remaining number of players. Thus, effectively, the comparison is a unit of gain for oneself and a unit of loss for the

[13] Their estimations are also in line with most of the previous estimations, please see the paper for details.

[14] More generally, any social preferences model that assumes the marginal utility of own earnings to be larger than the marginal disutility of earning more than another person weighted by the number of players would make this prediction. This includes models that include efficiency as a concern since in our baseline game, the Medium treatment, and in all the treatments in the second and third set of experiments, efficiency does not play a role.

victim. Therefore, all subjects are expected to prefer being part of the mob to their default payoffs.

4. The second set of experiments: The odd one out

A plethora of studies suggests that people often become bullying victims because they are different. When asked, school bullying victims state being different in appearance or in other things as the most important reasons for being bullied (Erling and Hwang (2004), Frisén et al. (2008)). In the workplace, envy seems to be an important reason for bullying (Zapf and Einarsen (2010), and references therein). In the second set of experiments we capture such asymmetries. We use the same payoffs as in the Medium treatment of the first experiment[15], with the exception that we modify the default payoff of one player. By changing this payoff we not only make this player naturally stand out, but we also change the relative position of earnings. We design two new treatments: In the Rich treatment, all payoffs for one player are increased by eight tokens, in the Poor treatment all payoffs are reduced by eight tokens. The additional payoff from bullying remains the same in Rich, Poor, and Medium, so they remain comparable in this respect. Moreover, as in the Medium treatment, bullying is redistributive in both the Rich, and the Poor treatments. Table III summarises the possible payoff constellations.

Table III here

Figure II and Table IV show the frequency of bullying over periods and the average bullying frequency per treatment, in the same fashion as Figure I and Table II of the first set of experiments. Coordination occurs even faster in the asymmetric treatments than in the symmetric

[15] We use the Medium treatment as the baseline because this generated mobbing rates of about one half – this means the results leave most room for treatment effects in both directions.

case: The average number of periods before the first coordination on a victim were 2.5 in the Rich and 2.3 in the Poor treatments[16] (the difference between Rich and Medium is significant at p=0.038, and between Poor and Medium marginally significant at p=0.095, both according to two-sided exact Mann-Whitney tests). So focality matters. This finding is clearly corroborated when we look at who is most likely to become the victim in the asymmetric treatments. Figure III shows that in three quarters of all cases of successful mobbing the focal player is the victim. This is regardless of whether the focal player is richer or poorer than the others; the fractions are very similar in both treatments. The observed percentages are highly significantly different from 25 per cent (which would be the outcome of randomly choosing the victim) according to the two-sided Wilcoxon signed-rank test with corresponding p values of 0.004 and 0.009. Note that there is only one focal, but three non-focal players; this means that a focal player was indeed nine times more likely to become the victim.

Figure II here

Table IV here

While the fraction of focal victims is very similar in both treatments, there is some evidence that the relative payoff of the focal player affects mobbing frequencies. Overall rates are similar in Medium and Poor (47.3% versus 57.1%, not statistically significant), but if the focal player is richer, bullying rates rise to 74.2%. This is significantly higher than in the Medium treatment. Envy towards the richer player seems to play some role in the likelihood of the richer player becoming a victim, while pity towards the poor does not. Notice that, a prefer-

[16] Again, we include all groups that coordinated on the same victim at least three times.

ence for efficiency is not relevant in explaining our results since in both the Rich and the Poor treatments, total payoff of the group does not change with bullying.

Figure III here

5. The third set of experiments: Greed or fear?

The results of the first and second set of experiments leave an important question open: Do players engage in bullying because they want to pursue material gain from mobbing, or out of fear of becoming the victim if they do not? To shed light on this question, we designed three new treatments. In the first of the new treatments, the set-up is the same as in the Medium treatment of the first set, with one difference: One of the four players is designated as the "safe" player. We call this the Safe-4 treatment (the '4' stands for the number of players, which we will vary). If the other three choose him as the victim, all players still receive their default payoffs. Trying to mob the safe player is simply ineffective (think of the safe player as the kid who is so big and strong that even all others united cannot hurt him). For the safe player, the fear of becoming the victim is ruled out as an explanation for bullying behaviour. As an indirect effect, the presence of a safe player should also calm the non-safe players' fears. If the safe player is less likely to bully, they are immediately less likely to be bullied, since no action can be taken against them without the safe player (recall that each player has veto power against mobbing somebody else). So if fear plays a major role, we should observe less mobbing in the treatment with the safe player compared to the Medium treatment.

This manipulation is open to the objection that the introduction of a safe player also reduces the number of victim candidates from four to three. This might make the task of coordinating

on a victim easier, and this could work against the reduction of bullying through the introduction of the safe player. Hence, by comparing mobbing rates in Safe-4 and Medium we could identify the presence of a fear effect (if we observe a drop in bullying, we do so *despite* having made bullying easier), but not its absence. If we do not observe any reduction in mobbing we could not clearly infer that fear plays no role, because a possible fear effect could be neutralised by easier coordination. To prepare our test for both scenarios, we introduce another treatment in which we also include a safe player – hence partially removing fear –, but make bullying harder than in Medium. We do so by introducing a fifth player to the game. One of the five is a safe player, the other four are not (we refer to this treatment as Safe-5). This sets the number of victim candidates back to four. Coordination on a victim should thus no longer be easier than in Medium. In addition, the number of bullies required increases to four, adding another difficulty to the coordination task. Notice also that, because we keep the total payoffs constant within each treatment, mobbing in the Safe-5 treatment results in six extra points for bullies rather than eight points of the Medium and Safe-4 treatments. Hence, if we do not observe a drop in bullying rates *despite* having made it harder and less payoff beneficial, we can dismiss the hypothesis of bullying out of fear with some confidence.

Table V here

Table VI shows the frequency of successful mobbing in the Safe-4 and Safe-5 treatments. Figure IV depicts the percentage of groups that engage in mobbing and the average number of nominations over the periods. These figures do not include cases in which players coordinated on the safe players – since this is ineffective, we do not count it as mobbing. In neither of the new treatments, the introduction of the safe player has any effect in reducing the incidence of mobbing, not even in the Safe-5 treatment, where the effect might be assisted by a

greater difficulty to coordinate and lower monetary gains from mobbing. Mobbing rates are very similar, and statistically indistinguishable, in Medium, Safe-4 and Safe-5. Hence, we do not find evidence that fear of becoming the victim explains the high mobbing rates in our experiment.

Table VI here

Figure IV here

Finally, we analyse whether we can detect mobbing out of fear at the individual level. Although we do not observe a drop in mobbing frequencies through the safe players, they may still be less likely to *initiate* bullying. This question cannot be answered with the existing treatments. By design, any successful bullying in Safe-4 and Safe-5 must involve the safe player, since it requires all players except the victim to coordinate. To identify differences in bullying behaviour between safe players and others we design a new treatment, which we call Safe-5-3. In this setting there are five players, but only three nominations on the same player are sufficient to establish a bullying group. This way it is possible to have bullying groups that do not include the safe player, and we can detect whether the safe player is more or less often a part of a three-person mob than the non-safe players. Table VII shows the payoffs for the case without mobbing, and with bullying groups that consist of either three or four players.

Table VII here

In the new treatment bullying rates are very high at 76.9%, significantly higher than in both other safe treatments and the Medium treatment. Perhaps this is unsurprising, since it requires only three out of five players to be inclined to bullying in order to establish a mob. Notice that, as in the previous treatments of this set, the overall payoff of the group remains unchanged with or without bullying in the Safe-5-3 treatment. Thus, once three nominations are received on a victim, the fourth person's decision boils down to deciding whether to join the mob to get a share of the spoils or stay out. The victim's payoff is unaffected and is equal to zero if there are three or four nominations. This might also create a bandwagon effect in which three nominations quickly lead to four. If such an effect exists, we would expect to see a decrease in the number of three nominations and an increase in the number of four nominations in the Safe-5-3 treatment compared to the Safe-5 treatment. We thus compare the average number of rounds with three and four nominations in the Safe-5 and Safe-5-3 treatments. We can see in Table VIII that the averages are almost identical. Figure V depicts the percentage of groups with three or four nominations on the same person, and the average number of nominations. The two treatments are almost identical in all these respects. Hence, there is no indication of the required number of nominations affecting the behaviour of the group, nor a bandwagon effect for the fourth players joining in faster in the Safe-5-3.

Table VIII

Figure V

Lastly, to study whether safe players are more likely to be part of a mobbing group than others, we look at all rounds in which a mob consisting of exactly three players has formed. In the Safe-5-3, bullying groups with or without the safe player are possible. We observe that

the safe player is a member of a three-player mob in 82.4% of all instances[17]. This is even higher than the chance rate of 75%, though not significantly so (p=0.401 with the Wilcoxon signed-ranks test with the null hypothesis that the median of the distribution is 75).[18] There is further no indication that the safe players in the Safe-5-3 treatment behave any differently than in the Safe-5 treatment: The safe player is a member of a three-player group in almost identically as many times in the Safe-5 treatment (83.2%), and again not significantly different than chance. Hence, in none of our tests do we find any evidence that removing fear of being bullied from the game reduces the frequency of mobbing in any way. Such fear therefore does not seem to be a likely driver of the frequent bullying we observe in the previous treatments.

6. Conclusion

We introduce the mobbing game, the first experimental paradigm to study the endogenous formation of bullying groups. The eight treatments of our experiment allow important insights into the nature of mobbing. Our results show that spontaneous formation of such groups is easy, and surprisingly many subjects are inclined to engage in bullying, showing little concern for the consequences to the victim. Nominations of victims are frequent even if the gains from it are minimal, and individuals do not hesitate to pick a fellow group member who is already poorer than themselves. In this case, the victim's relative poverty only serves as a focal point and a coordination device. Although fear of becoming a victim might seem a

[17] As with the other averages reported in this paper, this calculation is based on averaging over groups. Thus, for each group we counted the number of periods a three-player mob formed and checked the percentage of the times those groups included the safe player. We then averaged over all groups. For the Safe-5-3 treatment, there are 17 such groups and the standard deviation is 27.0. For the Safe-5 treatment, there are 14 such groups, the average is 83.2 and the standard deviation is 27.0.

[18] There are 20 possibilities to randomly form a group of three bullies and one victim out of five players. In 12 of them the safe player is part of the bullying group. In four of the remaining eight the safe player is the victim, thus these groups do not count as bullying groups. Consequently, by chance the probability to randomly draw a group with the safe player is 12/16=75%.

likely explanation for such ‘aggressive’ behaviour, our experiments do not produce any evidence for this.

Our results underline the importance of studying the group dynamics that lead to bullying, especially keeping in mind that our experimental subjects are from a very conventional subject pool, not one that would be naturally driven towards aggressive behaviour. Of course, our findings are not the last word in the matter. In this study we can touch only a few aspects of the strategic environment underlying bullying. However, our model can easily be extended to capture other potentially important variables as well, or to be applied to more specific subject pools.

References

Abbink K, de Haan T (2013): Trust on the Brink of Armageddon: The First-Strike Game. Working paper, University of Amsterdam.

Auf der Heide J (Dir.) (2009). *Van Diemen's Land.* Madman Entertainment. Film.

Borowsky IW, Taliaferro LA, McMorris BJ (2013): Suicidal thinking and behavior among youth involved in verbal and social bullying: risk and protective factors. *Journal of Adolescent Health*, S4-S12.

Cason T, Mui V-L (2007): Communication and coordination in the laboratory collective resistance game, Experimental Economics, 10, 251.

Centers for Disease Control and Prevention (2012): Youth risk behavior surveillance United States, 2011. MMWR, Surveillance Summaries 2012;61 (no. SS-4). www.cdc.gov/mmwr/pdf/ss/ss6104.pdf.

Cooper R, DeJong D, Forsythe R, Ross T (1992): Selection Criteria in Coordination Games. *American Economic Review* 80, 218-233

Cooper R, DeJong D, Forsythe R, Ross T (1992): Communication in Coordination Games. *Quarterly Journal of Economics* 107, 739-771.

Devetag D, Ortmann A (2007): When and why? A critical survey on coordination failure in the laboratory. Experimental Economics 10, 331-344.

Erling A, Hwang P. (2004):Swedish 10-year-old children's perceptions and experiences of bullying. *Journal of School Violence* 3, 33-43.

Falk, A., and Szech, Nora (2013a): "Morals and Markets". *Science* **340** (6133): 707–711.

Falk, A., and Szech, Nora (2013b): "Organizations, Diffused Pivotality and Immoral Outcomes". Working paper.

Fehr, E., and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817-868.

Forsythe R, Horowitz J, Savin N, Sefton N (1994): Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6, 347–369.

Frisén A, Holmqvist K, Oscarsson D (2008): 13-year-olds' perception of bullying: Definitions, reasons for victimisation and experience of adults' response. *Educational Studies*, 34, 105-117.

Gladden, R.M., Vivolo-Kantor, A.M., Hamburger, M.E., and Lumpkin, C.D. (2014) *Bullying Surveillance Among Youths: Uniform Definitions for Public Health and Recommended Data Elements, Version 1.0.* Atlanta, GA; National Center for Injury Prevention and

Control, Centers for Disease Control and Prevention and U.S. Department of Education. Retrieved from http://www.cdc.gov/violenceprevention/pdf/bullying-definitions-final-a.pdf

Hansen N (2000): The Custom of the Sea. Hoboken:Wiley & Sons.

Hoffman E, McCabe K, Shachat K, Smith V (1994): Preferences, Property Rights, and Anonymity in Bargaining Games. *Games and Economic Behavior*, 7, 346-380.

Isoni A, Poulsen A, Sugden R, Tsutsui K (2013): Focal Points in Tacit Bargaining Situations: Experimental Evidence, *European Economic Review*, 2013, 59, 167-188

Kowalski RM, Limber SP (2013): Psychological, Physical, and Academic Correlates of Cyberbullying and Traditional Bullying. *Journal of Adolescent Health* 53, S13-S20

Olweus, D. (1993). *Bullying at school: What we know and what we can do*. Wiley-Blackwell.

Olweus, D. (1994). Bullying at school: basic facts and effects of a school based intervention program. *Journal of child psychology and psychiatry*, *35*(7), 1171-1190.

Poulsen A, Poulsen O, Bett Z (2013): How Salient is an Equal but Inefficient Outcome in a Coordination Situation? Some Experimental Evidence. Working paper, University of East Anglia.

Productivity Commission (2010): Performance Benchmarking of Australian Business Regulation: Occupational Health & Safety. Australian Government Research Report.

Van Huyck J, Battalio R, Beil R (1990): Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure, *American Economic Review*, 234-248.

Van Huyck, J, Battalio R, Beil R (1991): Strategic Uncertainty, Equilibrium Selection Principles, and Coordination Failure in Average Opinion Games. *Quarterly Journal of Economics*, 106, 885–910.

Yang, Y., Onderstal, S., and Schram, A. (2013). Inequity Aversion Revisited. Tinbergen Institute Working Paper.

Zapf D, Einarsen S (2010): Individual Antecedents of Bullying. In Einarsen S, Hoel H, Zapf D, Cooper C (Eds.): Bullying and harassment in the workplace: Developments in theory, research, and practice. Crc Press.

Zapf D, Escartín J, Einarsen S, Hoel H, Vartia M (2010): Empirical Findings on Prevalence and Risk Groups of Bullying in the Workplace. In Einarsen S, Hoel H, Zapf D, Cooper C (Eds.): Bullying and harassment in the workplace: Developments in theory, research, and practice. Crc Press.

Table I

Payoffs in the Symmetric Treatments

Situation	Treatment	Player 1	Player 2	Player 3	Player 4	Total
Default	All	24	24	24	24	96
		Victim	Player 2	Player 3	Player 4	
Bullying	High	0	48	48	48	144
	Medium	0	32	32	32	96
	Low	0	25	25	25	75

Table II

Bullying Rates in the First Set of Experiments

	Bullying rate (%)	**Pairwise comparisons***
High	86.0	
		p=0.010 (M-H)
N=10	(11.7)	
Medium	47.3	
		p=0.018 (L-M)
N=13	(37.0)	
Low	13.5	
N=10	(21.4)	

*Mann-Whitney exact test, 2-tailed.

Table III

Payoffs in the Asymmetric Treatments

Treatment	**Situation**	**Focal**	**player 2**	**player 3**	**player 4**
Poor	Default	16	24	24	24
	Bullying Focal	-8	32	32	32
	Bullying other	24	32	32	0
Rich	Default	32	24	24	24
	Bullying Focal	8	32	32	32
	Bullying other	40	32	32	0

Table IV

Bullying Rates in the Second Set of Experiments

	Bullying rate (%)	**Pairwise comparisons**
Medium	47.3	p=0.024 (M-R)
N=13	(37.0)	p=0.443 (M-P)
Rich	74.2	
N=13	(29.4)	p=0.188 (R-P)
Poor	57.1	
N=12	(32.4)	

*Mann-Whitney exact test, 2-tailed.

Table V

Payoffs in the Safe Player Treatments

Treatment	Situation	Safe Player	player 2	player 3	player 4	player 5
Safe-4	Default	24	24	24	24	
	Bullying	32	32	32	0	
Safe-5	Default	24	24	24	24	24
	Bullying	30	30	30	30	0

Table VI

	Bullying Rate(%)	Pairwise comparisons
Medium	47,3	p=0.907 (M-S4)
N=13	(37,0)	p=0.936 (M-S5)
Safe-4	50,8	
N=12	(43,9)	p=0.636 (S4-S5)
Safe-5	48,3	
N=15	(30,5)	

Mann-Whitney exact test, two-tailed.

Table VII

Payoffs in Safe-5-3

Treatment	Situation	Safe Player	player 2	player 3	player 4	player 5
Safe-5-3	Default	24	24	24	24	24
	Bullying (3 Nom)	32	32	32	24	0
	Bullying (4 Nom)	30	30	30	30	0

Table VIII

	Avg # of Rounds with 3 Nominations	**Avg # of Rounds with 4 Nominations**
Safe-5	4,8	9,7
N=15	(3,9)	(6,1)
Safe-5- 3	5,7	9,7
N=18	(4,4)	(6,0)
	p=0.298	p=0.482

One-tailed Mann-Whitney exact test.

Figure I

Percentage of Successful Bullying and Average Number of Nominations per Group

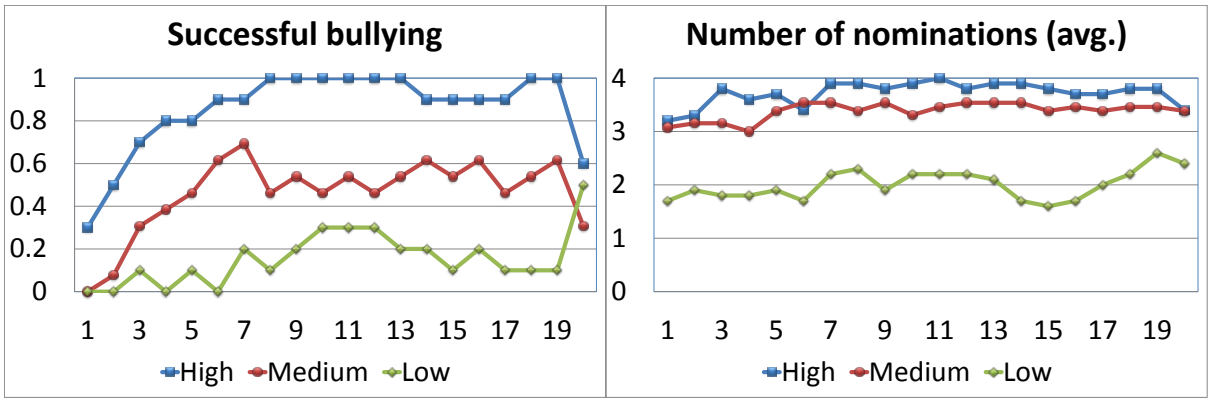

Figure II

Percentage of Successful Bullying and Average Number of Nominations per Group

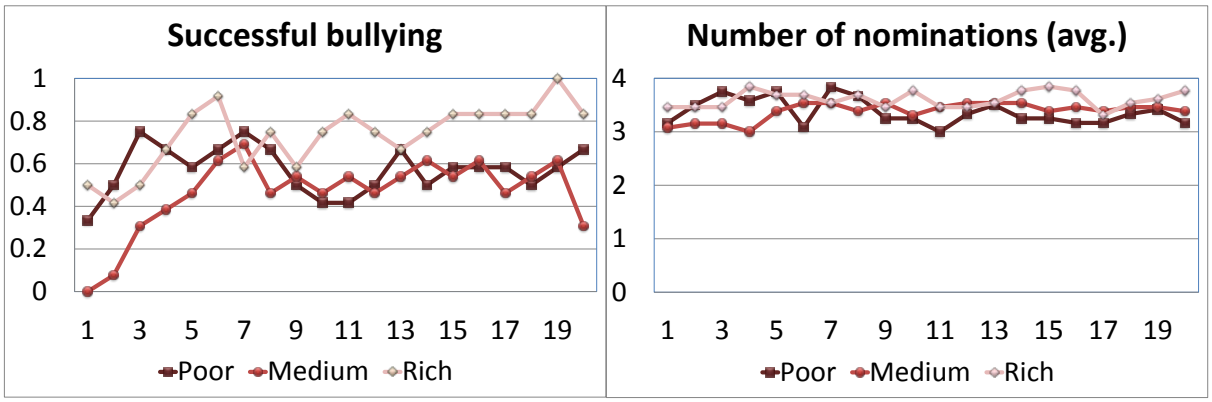

Figure III
Percentage of Victims that are Focal

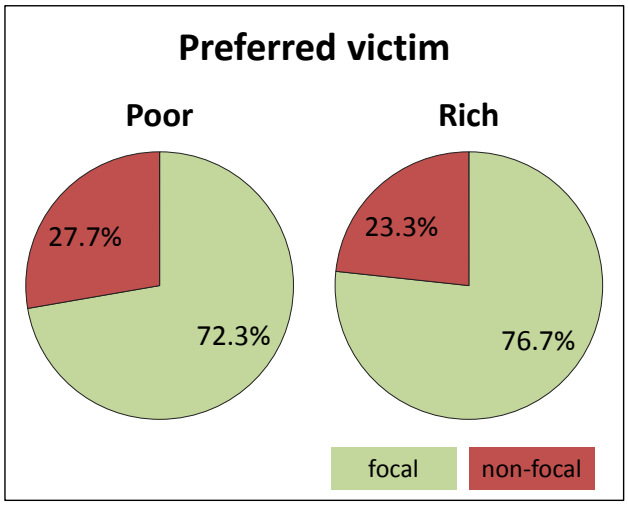

Figure IV

Percentage of Successful Bullying and Average Number of Nominations per Group

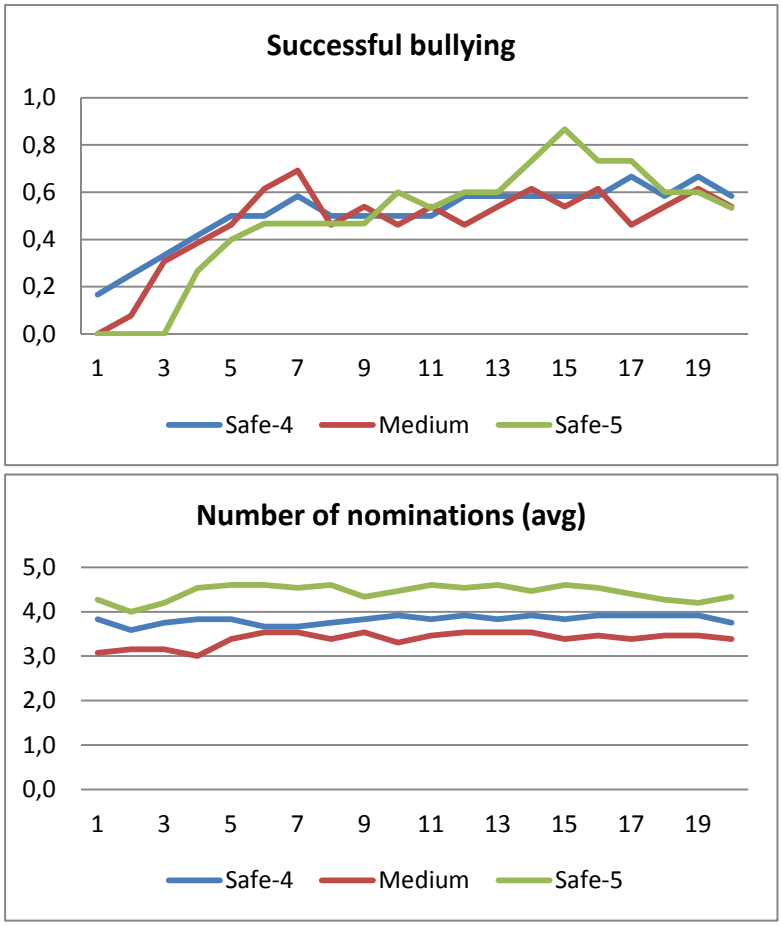

Figure V

Percentage of Groups with 3 or 4 Nominations on the Same Player and Average Number of Nominations per Group