


Altruistic Punishment in the Classroom: An Update on the Tragedy of the Commons Extra Credit Question

Teaching of Psychology
2019, Vol. 46(2) 153-157
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0098628319834208
journals.sagepub.com/home/top


Dylan Selterman¹

Abstract

A world-famous classroom exercise gives students the choice between 2 points (the communal option) or 6 points (the over-consumption option) of extra credit toward an assignment in their course, but if more than 10% choose 6 points, no one receives any points. In the current variation, students ($N = 795$) were also given a third option—to sacrifice their own potential points to take away points from another randomly selected student who chose 6 points. Across seven course sections, 19 students chose this option based on the concept of “altruistic punishment,” with many expressing concern about the entire class losing points. Most students had a positive attitude toward the exercise, viewing that it helped them understand new material and increased their interest in psychology, believing that psychology can help solve social problems. This exercise effectively demonstrates group cooperation dilemmas.

Keywords

altruism, dilemmas, social groups, social psychology, game theory

Behavioral scientists and game theorists have been using laboratory experiments to illustrate social dilemmas for decades (Camerer, 2003). Psychology instructors use similar exercises to illustrate those mechanisms in classroom settings. For example, Peden, Keniston, and Burke (1990) employed a “tough choices” game for students to earn extra credit after exams. In one version, students choose between a lower point option (e.g., 1 bonus point) and a higher point option (e.g., 5 bonus points), but if the total number of students choosing the higher point option exceeds a given threshold (e.g., 20%), then no student will receive any points. If the class stays under the threshold, they will receive the point total they indicated on their paper. I employed a similar version in my introductory psychology courses, in which I gave students the opportunity to choose 2 or 6 bonus points on their grade, but if more than 10% chose 6, no one would receive points. This illustrates social dilemmas involving the behavior of groups, including the *tragedy of the commons* (Hardin, 1968), which explains how people might be tempted to behave selfishly despite negative repercussions for their group as a whole. This is also referred to as the *n*-person prisoner’s dilemma (Bonacich, Shure, Kahan, & Meeker, 1976). This question became world-famous when a student tweeted about the exercise, which resulted in international news coverage. The exercise has resonated with millions of people worldwide (see Selterman, 2015a, 2015b).

The purpose of the current article is to illustrate a new variation of this classroom exercise to include another variable, currently referred to as *altruistic punishment*, an idea pioneered

to help explain cooperation in large, genetically unrelated social groups (Fehr & Gächter, 2002). This was originally demonstrated experimentally using a “public goods” game, in which people in small groups have the opportunity to contribute to a common fund (which will earn interest). However, the game is set up such that it would be economically advantageous for each individual to withhold contributions and hope for others to contribute their funds, which in theory would yield greater earnings for the withholding player relative to the contributors. But the maximum return for the whole group can be reached only if every player contributes. In the standard version of this game, cooperation is low and also decreases over time, relative to an alternate version in which players can also elect to “punish” those who do not contribute. This is costly to both the punished and the punisher—the punisher sacrifices their own currency in order to take away the currency of another player. However, contributions go up (conversely, social loafing goes down) when the possibility of altruistic punishment is introduced into the game. Thus, this mechanism is advantageous for groups to cooperate and thrive economically.

¹ Department of Psychology, University of Maryland, College Park, MD, USA

Corresponding Author:

Dylan Selterman, Department of Psychology, University of Maryland, 1147
Biology/Psychology Building, College Park, MD 20742, USA.
Email: dylan.selterman@gmail.com

The current article presents a new classroom exercise that merges the concept of altruistic punishment with the n -person social dilemma game described earlier. In this new version of the class exercise, students are again given a choice between a lower point option (2 points) and a higher option (6 points), with the stipulation that if more than 10% of the class chooses 6, then no one will receive any points. If the class (as a whole) stays below the 10% threshold, then each student will receive the point total they indicated on their papers. However, students are now also given a third option: They can sacrifice their possible points (i.e., choose 0 points), and if they choose this option, their points would go toward punishing another student (picked at random) who had chosen 6 points. This could, in effect, reduce the total number of 6-point choosers, such that the class would have a better chance of staying under the 10% threshold. Thus, choosing 0 points represents the *altruistic punishment* option—the student who chooses 0 points foregoes their own resources in order to help the group. Under this system, if the group manages to stay under the 10% threshold, then for each student who chose 0 points, one of the 6-point choosers would be selected at random to lose all of their points, but other 6-point choosers may wind up keeping their points (depending on the number of 0-point choosers in the class).

Method

Starting in the Spring 2016 semester and continuing through the Fall 2016 term, students in introductory psychology courses completed this exercise during a regular class day (not during an exam or other assessment). In total, 795 students participated across 6 course sections, $M_{\text{age}} = 19.17$, $SD = 1.37$, including 422 women and 241 men (gender data were not available for one section of 131 students due to human error). Additionally, a seventh section of 29 students completed this exercise during the summer, but follow-up data were not collected due to the fact that it was a class of under-18 high school students. The exercise is carried out during the logically appropriate unit of material in the course. In the introductory psychology course, the exercise is done during the unit on social processes, whereas in the social psychology course, the exercise is done during the unit on group processes. In both courses, this happens during the latter third of the course, at the point where students have had time to get to know one another voluntarily (though there were no activities designed to increase group cohesion before the exercise).

At the beginning of the exercise, students are told the rules of the game, and there is an opportunity for questions and clarifications. Then, students privately write down their answer choice underneath their name on a small piece of paper, fold it over, and hand it in to the instructor/teaching assistant (TA). The instructor and TAs then proceed to count the total number of 0-, 2-, and 6-point choosers. The students are given the overall results during class and then a discussion ensues. Students comment on the exercise, including emotional reactions and their thoughts about the psychological nature of groups. Students are invited to share their choice with the class and

explain the rationale behind their decision. As noted before, the procedure is meant to be confidential, such that students write their responses on secret ballot paper that no other students can see (the TA/instructor would need to see names in order to assign the extra credit). However, some students voluntarily choose to “out” themselves during the ensuing classroom discussion.

After discussion, the class hears a brief lecture on social dilemmas and group behavior, including the *tragedy of the commons* phenomenon, the prisoner’s dilemma, and game theory principles. This includes a presentation of various strategies to reduce overconsumption of resources, including social norms, behaviorist strategies (reinforcement), and the concept of *altruistic punishment*. Questions are answered as necessary. Finally, students complete a questionnaire that assesses their attitudes toward the exercise, their attitudes toward psychology in general, and to once again indicate their choice on the exercise itself. The questionnaires are confidential with no names/identifying information provided, but they are linked to research consent forms through a nonidentifiable code number. The consent forms and questionnaires are given to students with the incentive of extra credit toward their course grade if they choose to participate.

The questionnaire contained 7 items borrowed from Peden et al. (1990), including attitude items such as “I am very glad to have been involved in this activity,” and “This activity increased my interest in psychology.” In addition, new items were included to further probe whether students believed they were behaving in a prosocial or selfish way, as well as whether personal definitions of greed had changed as a result of the exercise. All were administered using a 7-point Likert-type scale. The full items are presented in Table 1, along with factor analysis results (see Questionnaire Results subsection).

As a learning outcome measure, I created 3 multiple choice content knowledge acquisition items that appeared for students on my final exam. These items tested students’ knowledge of what the tragedy of the commons is (on a definitional level and how it manifests in real-world situations), as well as knowledge about altruistic punishment, and how to effectively reduce overconsumption in large groups. I report data on 263 introductory psychology students who took the final exam, testing those who chose the correct answer against the distractors.

Results and Discussion

This section will be divided into five subsections. The first will focus on the results of the exercise for each section. The second will focus on the results of the questionnaire data. The third will contain more qualitative observations about students’ reactions. The fourth will include a report on learning outcome data. The last section will contain general commentary and conclusions.

Exercise Results

Table 2 displays the results for each section that completed this exercise, including the total number of students who selected

Table 1. Factor Loadings for the Exploratory Factor Analysis.

Item	Understanding	Positive Feelings	Prosocial Motivation
Making and discussing these tough choices promoted my knowledge of psychological science	.76		
Making and discussing these tough choices promoted my personal development	.67		
This activity increased my interest in psychology	.59		
My definition of “greed” has changed as a result of this activity	.49		
The purpose of this activity became clearer to me during our class discussions	.46		
I believe that psychology research can be used to help overcome social problems	.42		
It was not fair for our extra credit score to be influenced by the points that we received (or did not receive) for participating in this activity (R-scored)		.68	
I recommend using this activity as part of this class in the future		.61	
I am very glad to have been involved in this activity		.61	
I felt coerced or forced to participate in this activity (R-scored)		.55	
I believe that my classmates behaved selfishly during this exercise (R-scored)		.28	
I was motivated to “do right” by my fellow classmates			.71
I wanted to be altruistic/pro-social			.61
I believe that I behaved selfishly during this exercise (R-scored)			.55
I did not give much thought about whether my classmates would receive any bonus points (R-scored)			.50
I was concerned that if too many people chose the 6-point option, we would lose all points			.45
I wanted to be a hero			.34
I wanted the greatest number of points possible to me (R-scored)			.31
I would make a different choice if my class had the opportunity to do this activity again (R-scored)			.24

Note. Factor loadings reflect an exploratory factor analysis with Varimax rotation. Principal components analysis and Oblimin rotation yielded nearly identical factor structures.

Table 2. Exercise Results Per Course Section.

Semester Year	Option			Effective 6 Points
	0 Points	2 Points	6 Points	
1. Spring 16 (127)	2 (1.6%)	117 (92.1)	8 (6.3%)	6 (4.7%) ^a
2. Spring 16 (112)	5 (4.5%)	88 (78.6%)	19 (17%)	14 (12.5%)
3. Spring 16 (183)	3 (1.6%)	154 (84.2%)	26 (14.2%)	23 (12.6%)
4. Summer 16 (29)	2 (6.9%)	23 (79.3%)	4 (13.8%) ^a	2 (6.9%) ^a
5. Fall 16 (91)	2 (2.2%)	81 (89%)	8 (8.8%) ^a	6 (6.6%) ^a
6. Fall 16 (38)	0 (0.0%)	36 (94.7%)	2 (5.3%) ^a	2 (5.3%) ^a
7. Fall 16 (244)	5 (2.0%)	187 (76.6%)	52 (21.3%)	47 (19.3%)

Note. Each course section is listed in the left column with the total number of students in parentheses. For each option column, the total number of students who selected that option appears, with the percentage of the total number of students in parentheses. The final column represents the total number of people who received 6 points taking into account those students whose points were negated by the 0-point choosers.

^aDenotes that this course section stayed under the 10% threshold.

each of the three possible choices, along with percentages. As can be seen in the table, 4 of the 7 class sections successfully completed the task, meaning that they (as a group) effectively stayed under the 10% threshold of 6-point choosers, and received the points they selected. The other three course sections had more than 10% of the class choosing 6 points.

Questionnaire Results

To make analyses more efficient and reduce the likelihood of a Type 1 error, I ran an exploratory factor analysis on the

questionnaire items, which were reduced to three conceptually meaningful factors. The first factor, which I labeled *Understanding*, contained items relevant to personal growth and interest in psychology (“My definition of greed has changed as a result of this activity;” “I believe that psychology research can be used to help overcome social problems”). The second factor, which I labeled *Positive Feelings*, contained items that captured students’ attitude toward the exercise (“I recommend using this activity as part of this class in the future;” “I felt coerced or forced to participate in this activity” [reverse-scored]). The third factor, which I labeled *Prosocial Motivation*, contained items relevant to students’ prosocial tendencies while completing the exercise (“I was motivated to ‘do right’ by my fellow classmates;” “I did not give much thought about whether my classmates would receive bonus points” [reverse-scored]). One item was removed due to low loadings on multiple factors—this item was meant to capture whether students had learned about this exercise previously (in another course or through the media) and that influenced their response. Table 2 displays factor loadings.

The means for all three factors were significantly above the scale midpoint (4). Means, standard deviations, and test statistics for each factor appear in Table 3. Understanding correlated positively with Positive Feelings, $r = .33, p < .001$; and Prosocial Motivation, $r = .16, p < .001$; but Positive Feelings and Prosocial Motivation were uncorrelated, $r = -.004, p = .903$. A multiple regression analysis revealed that Positive Feelings and Prosocial Motivation both independently contributed to Understanding ($\beta = .33, p < .001, 95\%$ confidence interval CI [.24, .36]) and ($\beta = .16, p < .001, 95\%$ CI [.09, .22],

Table 3. Means, Standard Deviations, and Test Statistics for the Student Feedback Questionnaire.

Outcome Variable	M (SD)	t(794)	d	95% CI of Mean Difference
Understanding	4.66 (.92)	20.35***	.72	[0.60, 0.72]
Positive Feelings	5.17 (.99)	33.35***	1.18	[1.10, 1.24]
Prosocial Motivation	4.58 (.94)	17.46***	.62	[0.52, 0.65]

Note. $N = 795$. The comparison value for each t test was the scale midpoint (4). *** $p < .001$.

respectively), though Positive Feelings was a stronger contributor to the variance in Understanding compared to Prosocial Motivation. The total model explained 13% of the variance in Understanding, $F(2, 792) = 60.71, p < .001, R^2 = .133$.

A one-way analysis of variance revealed a main effect of exercise choice (0, 2, or 6) on Prosocial Motivation, $F(2, 792) = 245.408, p < .001$, whereas there was not a significant effect for Understanding, $F(2, 792) = 1.08, p = .340$, or Positive Feelings, $F(2, 792) = .87, p = .419$. Unsurprisingly, post hoc Bonferroni-corrected analyses revealed that the 6-point group ($M = 3.17, SD = .69$) scored lower in Prosocial Motivation compared to the 0-point group ($M = 5.01, SD = 1.04, d = -2.08, 95\% \text{ CI } [-2.31, -1.38]$) and the 2-point group ($M = 4.82, SD = .74, d = -2.31, 95\% \text{ CI } [-1.82, -1.47]$).

Qualitative Observations

Many students appeared frustrated or exacerbated immediately following the exercise if their class failed to stay under the 10% threshold, and many students attributed this failure to the “selfishness” of those classmates who chose 6 points. Conversely, students were elated and cheered if they achieved the objective and stayed under the 10% threshold. As noted above, during the class discussions following the exercise, many students voluntarily chose to “out” themselves by revealing their choices to the rest of the class, along with an explanation for why they made that decision. This included some who chose 6 points and some who chose 0 points. Some of the 6-point choosers noted that nobody had actually lost anything—specifically, that everyone’s grade was the same as what it was before, and that their decision was a low-risk, high-reward choice. Others spoke up similarly, noting that 2 points would not have contributed that much to their course grade, whereas 6 points could make a very real difference. The 0-point choosers, on three separate occasions, received applause and verbal cheers from the rest of the class upon revealing their choice. However, when describing their rationale, the 0-point choosers were not uniform in their responses. Some indicated that they wanted to be helpful to the group, while some indicated that the points were not very important to them and they did not consider it a big sacrifice. Nevertheless, it was clear that many students genuinely appreciated the personal sacrifice from those students who chose 0 points.

Learning Outcome

On the final exam for my introductory psychology course, I tested students on their knowledge about the tragedy of the commons, overconsumption, and altruistic punishment with 3 multiple-choice items. I used nonparametric one-sample χ^2 tests to compare expected values on the correct and distractor answer choices versus the observed distribution of choices. A majority of the class ($N = 263$) chose the correct answer for all 3 items [55%, 77%, and 62%], and choosing the correct answer (compared to the distractors) was unlikely to have occurred by chance, $\chi^2(4) = 248.08, p < .001$; $\chi^2(4) = 545.50, p < .001$; and $\chi^2(4) = 301.59, p < .001$. As a class, the students demonstrated a good understanding of these concepts. Similar results emerged for students in my social psychology course, who completed the same items on their final exam. A majority of this class ($N = 134$) chose the correct answer for all 3 items (73%, 91%, and 67%), and this was unlikely to have occurred by chance, $\chi^2(4) = 246.52, p < .001$; $\chi^2(4) = 423.99, p < .001$; and $\chi^2(4) = 194.67, p < .001$. Because I did not employ an experimental design, it cannot be concluded that participating in the exercise (and associated classroom discussion) necessarily caused a good understanding of those concepts as measured by exam performance. An alternative explanation is that students were able to learn these concepts independently from doing the exercise. Although this cannot be ruled out, it is worth noting that students reported that doing the exercise (and discussion) directly contributed to their knowledge of these psychological concepts, along with knowledge and interest in group cooperation and social processes. The learning outcome data merely complement the existing self-reported data from students.

General Conclusion

Before 2016, I had used the original version of the exercise (without the 0-point *altruistic punishment* option) in over two dozen psychology courses over 8 years, and virtually none of the classes stayed under the 10% threshold (just one section managed to stay under the threshold and received the points). Although no comparative data are available during this time period, it is abundantly clear that since the introduction of the 0-point option in 2016, group cooperation has increased overall. More than 50% of classes, in fact, stayed under the 10% threshold when combining the Spring, Summer, and Fall 2016 terms. This shows a trend (though not experimental evidence) for the effectiveness of altruistic punishment options to increase cooperation in groups of students. Future research might utilize other game theory variables to teach students about group cooperation and overconsumption.

The classes who successfully stayed under the 10% threshold for 6-point choosers were smaller in total size relative to those classes with more than 10% choosing 6. This is consistent with research showing that it is more difficult to get large groups of strangers to cooperate with

each other, though much of those experiments rely on groups smaller than 100 individuals (Nosenzo, Quercia, & Sefton, 2015). Furthermore, these studies show that when people realize the social benefits of cooperation, then cooperation increases even in large groups. It may be the case that the smaller classes feel more intimate and students feel more socially connected to their classmates, and this may have influenced their responses to some degree, but it is not clear that this influenced the decision to choose 0 points.

A common critique that was made during the viral fame of this exercise in 2015 was that it was “cruel” and that teachers who use it are subjecting students to a painful experience. Headlines included phrases such as “Is this the cruelest extra credit question ever?” However, this claim is not supported by the data, which show that students strongly enjoyed the experience, were glad to have participated in it, recommend it for future classes, and that it increased their knowledge/interest in psychology. Thus, it is best to temper criticisms of malevolence toward teachers employing these (or similar) exercises. There is evidence that students see inherent value in this activity and view it positively.

The question remains whether this type of behavior demonstrated in this exercise (specifically, choosing 0 points) truly reveals a purely altruistic, completely unselfish motive, and whether it generalizes to other settings outside of the classroom or laboratory. Many scholars continue to debate whether *pure altruism* exists at all, and others continue to debate whether altruistic punishment manifests in everyday, real-world behavior. Others are interested in theoretical implications of altruistic punishment as evidence in support of *multilevel selection* or *group-level selection* in evolutionary psychology/biology. I caution against using the results in this article as evidence for or against these ideas. The term *altruistic punishment* implies that people are selflessly sacrificing their resources (in this case, potential extra credit points) to facilitate their group. However, the data presented here from this classroom exercise do not reveal conclusively that students were 100% prosocial. Apathy toward extra credit and/or course could be an alternative explanation that we did not rule out, and there may be other mechanisms motivating people to choose 0 points. Similarly, one could make the argument that both 2-point and 6-point choosers were being selfish, each choosing the option that they felt would maximize the likelihood of them getting any points (2) or the most possible points (6). Furthermore, I also caution against overgeneralizing comparisons between students based on their decisions during this exercise. This is because the sample sizes for those who chose 0 and 6 points were small compared to those who chose 2, making empirical comparisons difficult due to a lack of statistical power. There was some evidence in the present data set that those who chose 6 points were less prosocially motivated compared to others,

but given the skewed group sizes, generalizability is limited. Future research should further explore this possibility.

As stated above, the primary purpose of this article is to demonstrate how the prisoner’s/commons dilemma game can also be utilized to effectively illustrate the concept of altruistic punishment. It is clear that students found the exercise rewarding and beneficial, and the introduction of the 0-point option appears to have increased students’ cooperation in terms of successfully staying below the 10% threshold, relative to previous years in which the 0-point option was not available.

Acknowledgments

Special thanks to Elham Khazaee-Torshizi, Katy Day, Lucy Venuti, and Romulus Castelo for their assistance with this project.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

References

- Bonacich, P., Shure, G. H., Kahan, J. P., & Meeker, R. J. (1976). Cooperation and group size in the N-person Prisoner’s Dilemma. *Journal of Conflict Resolution, 20*, 687–706.
- Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. New York, NY: Russell Sage Foundation.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*, 137–140. doi:10.1038/415137a
- Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 1243–1248.
- Nosenzo, D., Quercia, S., & Sefton, M. (2015). Cooperation in small groups: The effect of group size. *Experimental Economics, 18*, 4–14. doi:10.1007/s10683-013-9382-8
- Peden, B. F., Keniston, A. H., & Burke, D. T. (1990). Learning about individual and collective decisions: All for one and none for all. *Teaching of Psychology, 17*, 235–238. doi:10.1207/s15328023_top1704_5
- Selتمان, D. (2015a, July 20). Why I give my students a “tragedy of the commons” extra credit challenge. *The Washington Post*. Retrieved from <https://www.washingtonpost.com/posteverything/wp/2015/07/20/why-i-give-my-students-a-tragedy-of-the-commons-extra-credit-challenge/>
- Selتمان, D. (2015b, August 26). How understanding the prisoner’s dilemma can help bridge liberal and conservative differences. *The Conversation*. Retrieved from <http://theconversation.com/how-understanding-the-prisoners-dilemma-can-help-bridge-liberal-and-conservative-differences-46166>