Universidad del Desarrollo

USING EXPERIMENTAL GAME THEORY TO MEASURE COOPERATIVE RELATIONS IN ELEMENTARY SCHOOL CLASSROOMS TO UNDERSTAND ITS RELATIONSHIP WITH ACADEMIC PERFORMANCE AND SCHOOL CLIMATE.

BY: VICTOR LANDAETA-TORRES

A dissertation submitted to the faculty of government in candidacy for the degree of doctor of philosophy in social complexity science

THESIS SUPERVISOR:

DR. CARLOS RODRIGUEZ-SICKERT, UNIVERSIDAD DEL DESARROLLO

THESIS CO-SUPERVISORS:

DR. JORGE FABREGA, UNIVERSIDAD DEL DESARROLLO
DR. JORGE J. VARELA, UNIVERSIDAD DEL DESARROLLO
March, 2022
SANTIAGO, CHILE

A mis familiares y amigos, pasados, presentes y futuros.

ACKNOWLEDGMENTS

It's been a long, long way here. It's next to impossible to put on paper all the precise words to reflect how thankful I am for all of this process and to properly mention and honor all the wonderful people who have made this possible. A profound thank you to all of you, including everyone I would probably miss in this text, yet, here we go.

First of all, I would like to express my most profound appreciation to my supervisor, Carlos Rodriguez-Sickert. I am deeply grateful for your bet on me when I applied to the program, your patience and trust in me over all these years, your advice and wisdom, and your sincere friendship. This endeavor would not have been possible without Cristian Candia-Castro Vallejos. Thanks for your true friendship, encouragement, faith in me, advice, collaboration, and for all the funny and epic moments. I am very grateful to my co-supervisors, Jorge Fábrega and Jorge Varela, for your patience, comprehension, and guidance. I am also thankful to César A. Hidalgo, who gave me generous and smart advice, and kindly accepted me for a short stay as a visiting student at the MIT Media Lab. It was a wonderful experience in many different ways, and I am particularly grateful for getting to know amazing people from the Collective Learning Group. I'm incredibly grateful to all my colleagues who collaborated on the chapters of this thesis because it would have been impossible to reach this milestone without their hard work and generous disposition, and it was a pleasure to be working with: Melanie Oyarzún, Tamara Yaikin and Cecilia Monge. I am deeply indebted to Ana Taboada-Barber, leading the READ Lab from the College of Education of the University of Maryland, College Park, and all team members. Ana offered me the possibility of working on a research team after some time out of the academic world. She and the whole team made me feel motivated, valued, and capable of doing research again, and I am very thankful for that. It would have been way harder to find stamina and self-confidence for the last mile without this impulse.

I'm immensely thankful to all my classmates, colleagues, and professors at CICS-UDD; you were crucial in my training, this thesis work, and my well-being and happiness during all these years. I am especially grateful for the profound and insightful conversations, friendship, and funny moments with Tamara Yaikin, Cecilia Monge, Paz Santander, Jorge Castillo, Mauricio Aspé, Ricardo Guzmán, M. Ignacia Rivera, Claudia Fernandino, Paola Gómez, Francisca Martńez, Melanie Oyarzún, Gabriela Valdebenito, Yerka Freire, María Paz Martínez, Carina Sturla, Nicolás Ferrer, and Oriana Figueroa.

Many thanks to my friends from my physics undergraduate time; you were determinant for creating more hilarious anecdotes than I can remember and for being such great friends and persons. Love you: Christian Diaz, Manuel Alamo, Javier Cancino, Ignacio Alvizú, Isabella Liedtke, Pedro Salas, Morin Roa, Mauricio Sarabia, Rosario Astaburuaga, Mayra Cerda, Pablo Marchant.

Thanks also to my dear friends from my teaching experience in Limache that unforgettable year 2013 and how that ended up with me applying for this Ph.D.: Mónica Cerda, Rodrigo Sarmiento, Ignacio Lobos, Claudia Araya, Dinka Zúñiga and Camila Villanueva

Thanks to my beloved High School friends, who mandated that they must be on this page this time, given that they were not named in my physics B.Sc. final work: Javier Ramos, Alvaro Delgado, Hugo Navarrete, Nicolás Delpin, Roberto de la Paz, Alexandra González, and Gabriela Martínez.

Thanks to all our wonderful friends who supported and gave us great times and memories while living, working, and surviving through a pandemic far away from home in Maryland: Nitin Luthra, Ruchi Saini, Virginia Gomes, Juan Castillo, Araceli Alegría, María Paz Castillo, Pedro Roje, Florencia Sánchez, Russell King, Perry Melazzo, Megan Verrecchio

My deepest gratitude to my extended family: Love y'all, I wouldn't be here if it weren't for you. To my Grandparents Hugo, Nora, Sergio, and Elda; I'd like to think that you are proud of how educational levels go up in the family. To my uncles, aunts, and cousins: thank you for believing in me way more than I do. To my dad, Leonardo Landaeta: thank you for everything, for the conscious efforts for not asking about the thesis at some moments; I know it is hard for you, and for all the support during that strange but, in the end, fun trip to México. To my mom, Pamela Torres: I've always felt your unconditional love, support, and confidence in me; I admire and try to emulate some of your willingness and determination to everything, particularly to work; that also helped me force myself into the writing discipline. To my brother, Álvaro Landaeta: I feel grateful for being your brother and that we're part of each other's life. Thank you for calling me and reaching out to me, particularly this last period when I'm always happy to see you but rarely call you. To my sis-in-law, Paola Vallarino: I'm grateful and happy that I can call you family. To my godson, Juan Pablo: I love you nephew, and I hope that one day, many years later, you can read this and feel all my gratitude through these lines because you were an inspiration and a motivation not to give up, to be able to write you this message here. To my sister, Pamela Landaeta: we shared a whole lot of this process over these years; you entered this University too, and that brought us many extensive commuting together, which ended up in sightseeing remarkable places on the way, over and over again, and filling awkward silences with fascinating conversations. I also felt in a good way complied to finish because I have felt proud and inspired by your academic and personal growth over these years. I'm sure that if you go for a Ph.D. or whatever you propose, you'll rock it.

To my wife and life-partner, Francisca Vizcaya Jofré. I could not have undertaken this journey without you. And I wouldn't wanted either. So much has passed during this extended period: adopting our cat, moving together, getting married, you applying and being selected for a master's program, moving to the US, finishing your program, and moving back to Chile. I've always felt your support and love during this process, over my darkest moments when I was overwhelmed, hopeless, and tired of everything. Thank you so much for your deep comprehension over all this path and the painful and almost inscrutable inner process that I went through, and that at many moments you received the worst part. I really appreciate it, even when I don't say it enough. Thank you again to you, and thanks again to life for having this beautiful family with you and Gordo. I love you, and I want to keep walking by your side, as I wrote you in another acknowledgments section ten years ago.

0.1 Studies Included in this Document

Reciprocity heightens academic performance in elementary school students (Chapter 1)

Cristian Candia-Castro-Vallejos, Melanie A. Oyarzún-Wolf, **Victor Landaeta-Torres**, Tamara Yaikin, Cecilia Monge, César A. Hidalgo, Carlos Rodriguez-Sickert.

Cooperative relationships and bullying/victimization in the classroom: insights from an experimental approach (Chapter 2)

Victor Landaeta-Torres, C. Candia-Castro-Vallejos, Jorge Fábrega, Jorge J. Varela, Tamara Yaikin, Cecilia Monge, Carlos Rodriguez-Sickert.

Contents

0.1 Studies Included in this Document . . . 6

Abstract **9**

1 Reciprocity heightens academic performance in elementary school students **11**

1.1 Introduction . . . 12

1.1.1 Network theory and social capital . . . 14

1.1.2 Experimental game theory and cooperation . . . 15

1.1.3 Network-based measures of social capital and experimental game theory . . . 16

1.2 Methods . . . 17

1.2.1 Sample . . . 17

1.2.2 The Game . . . 17

1.2.3 Network measures . . . 19

1.3 Results . . . 20

1.4 Discussion . . . 26

Bibliography . . . 29

2 Cooperative relationships and bullying/victimization in the classroom: insights from an experimental approach **39**

2.1 Introduction . . . 40

2.2 Methods . . . 44

2.2.1 Participants (Sample) 44
2.2.2 Procedure 45
2.2.3 Instruments and Measures 45
2.3 Results 50
2.3.1 Descriptive Results 50
2.3.2 Multilevel Analysis 52
2.4 Discussion 58
Bibliography 61

Appendix A Chapter 1: Supplementary Material **67**
Bibliography 78

Appendix B Chapter 2: Supplementary Material **79**
Bibliography 84

Appendix List of Figures **85**

Appendix List of Tables **88**

Abstract

Cooperation has been key to our success as a species. Cooperative behavior on multiple levels and the structures that emerge from those interactions have been extending over different domains and scopes, from small food-sharing networks among hunter-gatherers to global trade and the generation and diffusion of knowledge and technology worldwide.

Groups and societies have been creating and changing increasingly complex social mechanisms and institutions to allow, promote, and support cooperation and its scaling. All these efforts have boosted and expanded the benefits of cooperation but also have increased our interdependence between different actors, which also implies higher systemic risks.

Cooperative Networks in small groups are the essential building blocks to sustain and promote large-scale cooperation. Moreover, large-scale cooperation is possibly our best chance to face today's most significant global challenges and threats, such as climate change and global pandemics.

Schools are crucial for human development. Here, we first face peers and learn how to cooperate with strangers. The school's main goals are socialization, the transmission of culture, and teaching knowledge and skills. Classrooms, in many ways, have a similar tribe-like structure and a nested and controlled general situation, so dynamic cooperative networks emerge and mutate. Two main challenges for schools nowadays are about learning and well-being. In specific, how can we improve learning and achievement? And, how can we make school a safer, bullying-free place?

Cooperative learning literature has consistently found that creating more cooperative learning environments has causal effects on increased academic achievement and decreased bullying behavior. However, no precise mechanism is known. So, our main research question over this work is the following: How cooperative networks in the classroom relate with other educational outcomes, particularly Academic Achievement and Bullying Behavior?

But, how can we measure cooperation in the class, and elicit these networks? We propose that we can use Experimental Game Theory, specifically implementing in the field an adaptation of a lab game-theoretic social dilemma, which try to reproduce with ecological similarities the tension

between individual interests and social efficiency that students face in real life everyday interactions in the classroom.

The essential concept is External Validity. We expect that the student's behavior in the game will reflect their day-to-day behavior and classroom structure. The dyadic social dilemma we implement is a modified version of the Prisoner's dilemma. A major adjustment is that in this game, students are aware of the partner's identity that has been matched to them each round, i.e., the game is non-anonymous. Thereby, their choices when playing again each other are not only the result of intrinsic prosocial dispositions (or their absence), but also the result of their history and the perceptions they have about each other.

In the first chapter, our first draft paper tackles the relationship between cooperative network topology and academic performance. We measured the cooperative centrality of students by quantifying their deviations from the average level of reciprocal cooperation in each interaction and found that students that engaged in high levels of reciprocal cooperation have significantly higher GPAs. In the second chapter, our second draft paper characterized the relationship between cooperative network topology and bullying subtypes. We categorized students on their bullying involvement in four different social categories: bully, bully-victim, victim, and non-involved students. Then, we use the data from the experiment combined with a self-report instrument and using multilevel modeling, we study how bullies, victims and bully-victims differ in their access to the elicited cooperative network. We found that bully/victims and victims tend to receive less tokens than non-involved students.

Both articles have important policy implications to the extent that they can inform the design of interventions in the early phases of education to improve both academic achievement and social coexistence.

Chapter 1

Reciprocity heightens academic performance in elementary school students

Abstract

Social relationships are pivotal for human beings. Yet, we still lack a complete understanding of the types and conditions of social relationships that facilitate learning among children. Here, we present the results of a study involving 855 elementary school children from 14 different public schools in Chile designed to understand their social learning strategies in classrooms. We mapped students' social relationships using a behavioral experiment–a non-anonymous social dilemma–that allows us to measure cooperation and infer reciprocal and asymmetrical relationships between peers. We implemented the experiment synchronously in each classroom using networked tablets and a friendly user interface to mitigate cognitive barriers and boost students' engagement. We found a positive and significant association between reciprocity and academic performance using regression models. This result holds after controlling for class attendance, sex, parents' education, social status, individual cooperative dispositions, and fixed effects per class group. Finally, using a difference-in-differences-like framework, we found more evidence that reciprocity seems to heighten academic performance by comparing two consecutive academic semesters. This effect is heterogeneous and is considerably more prominent for the top 20% of students experiencing higher levels of reciprocity in their social relationships. We expect these results to inform cooperative learning interventions in elementary education.

1.1 Introduction

"No significant learning can occur without a significant relationship". In this statement from a lecture given at the Education Service Center in Houston, Texas, in 1995, Dr. James Comer from Yale University clearly expresses the critical role of social relationships between students and their peers, teachers, friends, and family for learning. Certainly, this sentence does not mean that we cannot learn from people with no direct relationships; in fact, we can learn from unrelated people, books, and even online forums. However, knowledge and experiences acquired from meaningful relationships are remembered and applied more than others. Dr. Comer's quote also suggests that social relationships come first, and learning will follow. Here, we will take this insight and provide quantitative evidence on how reciprocity in social relationships is beneficial for learning in elementary school students.

In recent decades, several studies have shown a significant association between students' position in their social networks and their academic performance at all ages [1–13]. Academic performance correlates positively with social capital–the individuals' network of connections and tacit cooperation [14]–among college students in online degree programs [4], and with the flow of online and offline communication among undergraduates [8]. Yet, we still lack a complete understanding of the connection between social networks and learning, particularly among elementary school children.

Long-term returns to education depend mostly on early learning outcomes [15–17]; thus motivating the study of factors that contribute to the learning in elementary school students at both the individual and group levels. At the individual level, academic achievement has been shown to improve with parental education [18] and the involvement of parents in school activities [19]. Student time management [20], sleep quality and duration [21], physical activity [22] and, internet use [23], also contribute positively to learning. Here we attempt to shed light on how social relations within the classroom influence learning outcomes.

At the group level, positive externalities –peer effects– are known to be pivotal for social learning [24–28] and play a key role in academic outcomes [5; 8; 29–31]. Indeed, how people capture the effects of their social relationships within a particular social environment largely depends on the teaching strategy [10] and the social structure of their cooperative relationships [32]. For instance, at the dyadic level, social learning can be understood as a natural form of pedagogy, where cognitive mechanisms enable the transmission of cultural knowledge through imitation and communication [33]. What distinguishes this natural pedagogy from other types of social learning, e.g., prestige-biased social learning [34; 35], is that it not only requires the disposition to learn from the "student-role subject," but the willingness of the "teacher-role subject" to share their knowledge [33]. Thus, from a game-theoretical perspective, a pedagogic act between peers [36], qualifies as an act of cooperation in a standard social dilemma, a scenario in which individual and collective interests collide because there are incentives to maximize individuals' payoffs that generates a sub-optimal collective performance [37].

Capturing cooperative dispositions between individuals is challenging and often requires multi-dimensional instruments to uncover social relationships. Efforts to capture social networks date back to Jacob Moreno's sociograms [38]. Under Moreno's approach, the social network is obtained through surveys that ask students to state both whom they like or dislike to spend time with and who their friends are [6; 39–42]. Traditionally, relational studies are conducted through surveys. However, for primary school students, survey-based social network mapping may not be sufficient because of the exacerbation of different types of biases [43] such as the social desirability bias [44] (over-reporting of socially desirable behavior); cognitive barriers (difficult to establish that subjects fully understand the questions) [45]; and lack of engagement (length or unfriendliness of instruments generate poor answers) [46–49] associated with the implementation of self-report based instruments.

To tackle these biases, we implemented a game-theory-based experiment in which all the students of a given class play a dyadic social dilemma with each classmate. The game is played on networked tablets using a friendly drag and drop interface. This methodological approach facilitates the behavioral mapping of cooperative relationships by settling elementary school students in a familiar and ecological interactive environment. The advantage of using game theory to map the social network is twofold: first, due to the non-anonymous character of the game, it allows us to capture in a more comprehensive way the nature of cooperative relationships among students who, in most cases, have been together in the same class group for more than three years [50]; and second, the interactive nature of the game in which different actions lead to different payoffs mitigates the biases related to survey-based instruments [51].

With these measurements, we study the following question: Do students who participate in more mutually cooperative relationships increase their GPAs more than other students? To do this, we go through the following structure: The rest of this section reviews previous works on network science, social capital, and experimental game theory. The following section explains in more detail the game-theoretical experiment and its interface, data, and methods used. We then show the results and the econometric strategy to establish outcomes, and finally, we discuss the conclusions, limitations, and future work.

1.1.1 Network theory and social capital

Studies have found a positive relationship between the degree centrality and measures of individual social capital and the number of alters who may provide social support [52]. Egocentric network density moves beyond a simple count of the number of members in the network to consider the extent to which alters know one another. Thus, a respondent provides information on the relationships between the alters in their network. For a general account of network-based measures of social capital, see [53].

An alternative approach to measuring the resources provided by a particular link is to focus on the tie's strength, following sociologists' early work on social networks. In his seminal paper, Granovetter states that tie strength is a "... (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie" [54](p. 1361). Standard measures of tie strength include frequency and duration of contact, whether the tie is emotionally supportive, and whether it is multidimensional

[55].

Most of these studies rely on the survey-based elicitation of social networks. This approach in the context of school communities riddles back to the studies by Jacob Moreno [38]. Indeed, his 1934 book "Who Shall Survive?" contains some of the earliest graphical depictions of social networks (sociograms). Since then, social network analysis has been upgraded by techniques coming from graph theory such as alternative measures of centrality [56], and community detecting algorithms [57]. However, survey-based methods still prevail in current approaches that study the relationship between the topology of the social structure of the class and its link with individual academic performance [5].

1.1.2 Experimental game theory and cooperation

The implementation of social dilemmas that are characterized by the tension between individual and social incentives, such as trust games [58], public goods game [59], or common-pool resource games [60], provided empirical evidence for behavioral anomalies concerning the prediction of the homo economicus model (individuals' perfect rationality), understood as a simplified version of Stuart Mil's original abstraction. These deviations from the standard model have been interpreted as the result of other-regarding preferences [61] such as altruism, reciprocity, or inequity aversion [62].

The first wave of experimental studies mainly used Graduate students as experimental subjects, referred to by Henrich et al. [63] as WEIRD (Western, Educated, Industrialized, Rich, and Democratic) subjects and involved anonymous interaction. Then, the experimental literature has slowly moved to study natural populations [64] and involve games implemented in the field which stylistically represent social dilemmas that agents face in real-life [65; 66]. The new body of literature that studies the external validity of game-theoretic experiments states that cooperative social norms that are prevalent in the real world can penetrate laboratory behavior [67]. Individuals who are more cooperative in the real world also behave more cooperatively in the lab [68; 69];

and groups who achieve higher levels of cooperation in the real world also achieve higher levels of aggregate cooperation when playing a social dilemma in the lab [70–72]. The social domain of all of these experimental studies ranges from the fishers of Toyama Bay [70] through the exploitation of benthic resources in the Chilean coast [73], to the Wikipedians studied by [69]. These results suggest that individual contribution to social capital and group social capital in the real world can be measured in the lab using game-theoretic social dilemmas that reproduce the tension between individual interest and social efficiency that agents face in real life.

To the extent that one of the major threats to the external validity of economic experiments, as argued by [74; 75] refers to the anonymous character of interaction within the lab, we should expect more significant levels of external validity in our experiment. Moreover, a recent study that shows that non-anonymous interaction increases cooperation in contrast with anonymous interaction suggests that pre-existing social connections do affect laboratory behavior [50].

1.1.3 Network-based measures of social capital and experimental game theory

In the study of cooperation between humans, experimental game theory has contributed to understanding mechanisms that explain phenomena such as cooperation in collective action problems or the dynamics of reciprocal exchange [76]. Traditional research has mainly focused on anonymous interaction. However, group and dyadic social dilemmas in real life are often played between known subjects with previous and future history embedded in a social network structure. A school classroom provides an ideal environment for controlled social dilemma experiments whose results can be contrasted against real-life indicators of school life.

The possible outcomes for any dyadic game are: reciprocal mutual gains, indifference (or negative reciprocity), or abuse (one of the parties benefits at the expense of the other party) will be used as a proxy of the nature of the relationship in real life for such dyads.

Taking together that learning is a collective enterprise [24], and that several studies have shown the importance of peer effects on school performance (for a survey on these results, see [29]), we expect that the position of the individual in their social network and the strength of their ties could modulate the capacity of the student to capture learning externalities. Given that these externalities are generated by their peers during the learning process, the academic performance should be positively affected, as suggested by previous theoretical work [32].

1.2 Methods

1.2.1 Sample

We collected data from 855 students (between the 3rd and 5th grade) with an average age of 10.16 ± 1.18 (57.5% were females) in 14 different public schools and 45 classrooms, at the beginning of the second semester of 2017. We also collected administrative records for each student, including gender, educational level of student's parent or guardian, classroom attendance, and GPA. Most of the students in the sample have been members of the same classroom for more than three years, spending around eight hours together each school day. We also collected again academic performance and attendance in December 2017. Finally, we note that we have complete information for 769 students.

Our data collection methods and the experimental protocol were approved on May 4, 2016, by the human subjects review board of Universidad del Desarrollo (IT15I10079).

1.2.2 The Game

To measure relational cooperation, we implement a modified Prisoner's Dilemma using a friendly user interface on tablet computers (Fig. 1-1 A). Our design involves two modifications concerning the standard experimental design. First, the interaction is non-anonymous. In each round, students know who their counterparts are. Under standard game-theoretic experimental protocols, which involve anonymous interaction, networks elicited in the lab emerge from scratch, mainly through

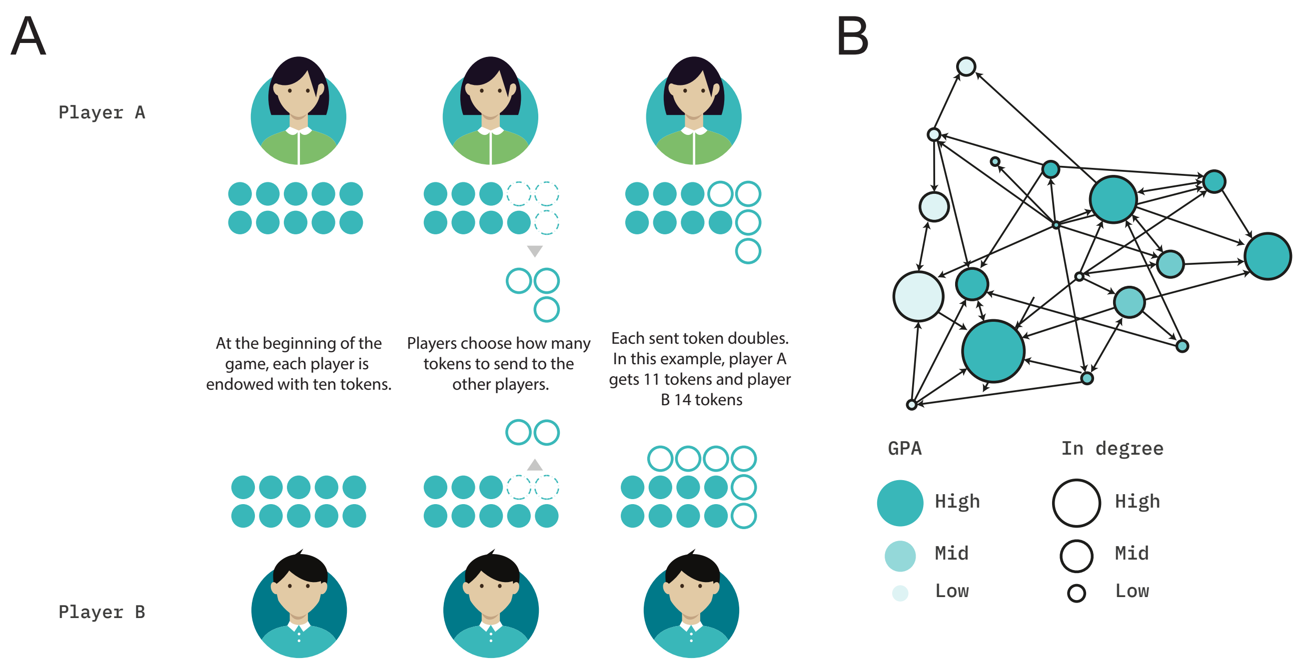

Figure 1-1: Experimental game. A) Students play a social dilemma; example of a dyadic interaction: i) Both players are endowed with 10 (ten) tokens. ii) Simultaneously, player A decides to send 3 (three) tokens, and player B decides to send 2 (two) tokens. iii) Sent tokens are doubled. iv) Player A receives 4 (four) tokens and player B receives 6 (six) tokens. B) A students' network for a single classroom. For the sake of visualization, edges correspond only to fully cooperative interactions.

assortative interaction between anonymous players [77]. However, to capture the nature of pre-existing relationships, we departed from the standard protocol and considered non-anonymous interactions. Thereby, we expect to capture not only the individual's general intrinsic disposition to cooperate with others but their dispositions to cooperate with each specific classmate, taking into account their particular context, perception, and history together [50].

Second, rather than deciding whether to cooperate, students can implement different levels of cooperation by sending a positive amount of tokens or choosing not to cooperate and keep all of their tokens. In each dyadic interaction, both students are endowed with ten tokens. Then, each student decides simultaneously how many tokens to give to the other students. Every received token doubles, so students get the tokens they keep plus twice the tokens they receive. Thus, we created a social dilemma where there is a tension between individuals' incentives (keep all tokens) and social incentives (give tokens that are doubled to their peers with whom they share a history and common past). Concretely, both players in a dyad would get more tokens if they give each other all their endowed tokens (receiving 20 tokens each) rather than keeping them (10 and 10). However, each of them has unilateral incentives to keep the endowed tokens because in every round, regardless of the classmate's decision, a player would always end up with more tokens in a round by not giving any tokens to that classmate, and vice versa.

The experiment ends once every student has played with all their classmates. We use these cooperative interactions to map students' networks. For instance, Figure 1-1 B shows a classroom network in which nodes represent students, and directed links indicate fully cooperative interactions (a student giving ten tokens to another student).

1.2.3 Network measures

We quantified individuals' cooperation and reciprocity in their classrooms using network measures. We define a weighted adjacency matrix for each classroom w_{ij}, representing the number of sent tokens from student i to student j. Table 1.1 shows the network metrics used in this study.

Network measure	Social Capital	Formula
Average in-degree	Average received cooperation	$r_i = \frac{1}{N}\sum_{j\neq i} w_{ji}$
Average out-degree	Average sent cooperation	$s_i = \frac{1}{N}\sum_{j\neq i} w_{ij}$
Reciprocated weight	Reciprocated cooperation	$R_i = \frac{1}{N}\sum_{j\neq i} min[w_{ji}, w_{ij}]$
Page-Rank	Social ranking	$Rank_i = \frac{1-d}{N} + d\sum_{j=1}^{n} \frac{w_{ij} Rank_j}{\sum_{k=1}^{n} w_{kj}}$

Table 1.1: Network measures. w_{ij} is the number of tokens sent from i to j, d represents a dumping factor ($d = 0.85$ following [78]), and N is the number of students in each classroom

Average received cooperation, r_i, measures the average cooperation received by ego (i). Average sent cooperation, s_i, measures the average sent cooperation. Reciprocated weight, R_i, measures the average level of reciprocity for each ego (i). $Rank_i$, calculated using Page-Rank [78], measures the relative social ranking of students based on the network of cooperation.

1.3 Results

Figure 1-2 shows the emerging patterns of token sendings among students, where the total number of sending combinations among classmates is $18,334$. Panel A shows the bivariate distribution of sent and received tokens. Peer interactions concentrate close to the corners. In fact, more than 15% of the exchanges are fully cooperative (Figure 1-2 A I), while 22.1% of interactions are highly defective and involve both students sending two or fewer tokens to each other (Figure 1-2 B II). Asymmetric interactions are also visible in the behavioral game. About 12% of interactions involve a student sending ten tokens and getting two or fewer tokens in return (Figure 1-2 B III). Also, we observed that each student sent tokens across the entire range of values (Supplementary Figures 3.2), suggesting that the previous history of interactions among a pair of students matters for their decision-making process [50] above and beyond pure strategies such as always cooperate or always defect.

Figure 1-2 B depicts a histogram summarizing the amount of sent tokens. The figure shows that despite students could send any number of tokens between zero and ten, they mainly engaged in either fully cooperative (sending ten tokens) or rather non-cooperative strategies (sending two or fewer tokens).

Now, we ask, do students who participate in a high number of mutually cooperative relationships increase their GPAs more than other students? We investigate if reciprocity, defined in each exchange as the minimum amount between sent and receiving tokens (See Table 1.1), plays a role in academic performance. Here, the ideal situation to correctly identify the reciprocity's causal effect in academic performance would be to exploit an exogenous variation in reciprocity. Still, it is impossible to create such a variation in this type of experimental design due to the intricacies between reciprocity and peer interactions. Instead, we rely on statistical tools [80] to estimate the individual future GPA as a function of the individual-level average reciprocated cooperation, controlling for different confounders. Omitted variables simultaneously determining reciprocated cooperation and GPA improvement or reverse causality could provide biased point estimations.

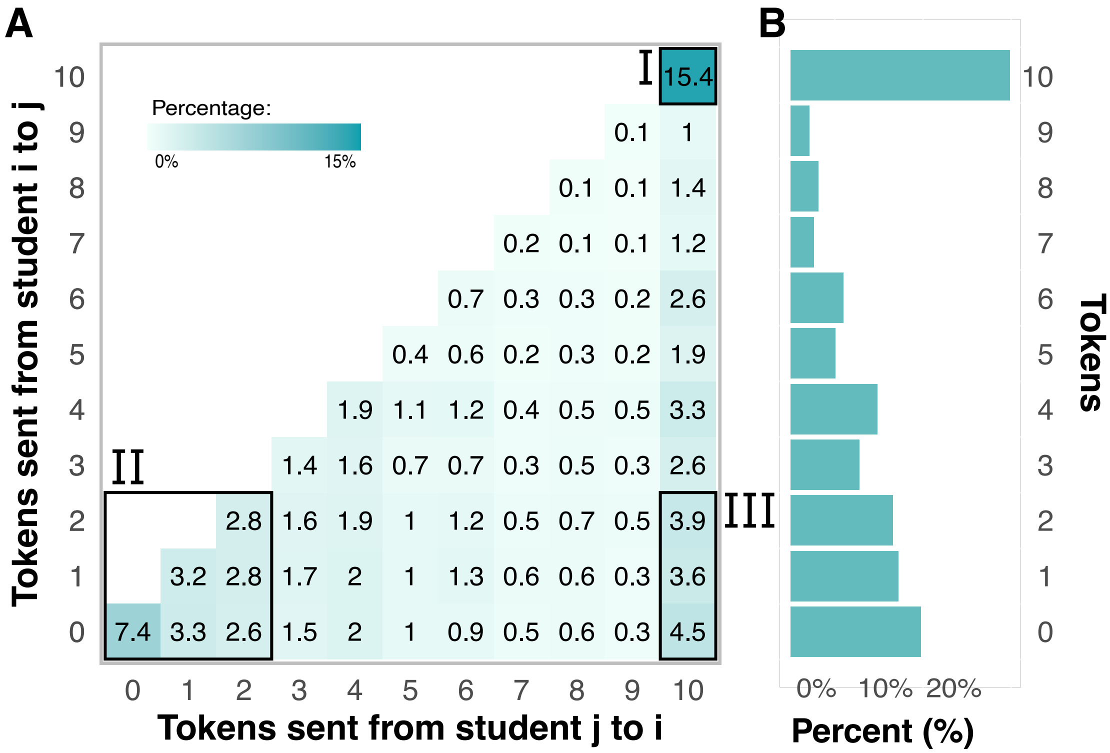

Figure 1-2: Patterns of sent and received tokens across all peer interactions. A) Bivariate distribution of sent tokens. We observe three peaks describing the most commonly used strategies: I) Social optimum (top-right). II) Nash equilibrium (bottom-left). III) Asymmetric exchange (bottom-right). B) Univariate distribution for sent tokens.

Remarkable possible omitted variables affecting both academic performance and reciprocity are intelligence, illness, and socioeconomic background (See Supplementary Notes 1.1 for more detail); therefore, we provide proxy variables for them–prior academic performance, attendance, and guardian's education, respectively. However, we acknowledge that other unconsidered omitted variables could play a role in the identification process. Yet, our statistical tools help us unveil a cleaner effect of reciprocity on future GPA.

We rely on a two-fold identification strategy: first, refining the statistical model with relevant controls and a fixed effect controlling for average classroom characteristics. Second, we implement a complementary treatment intensity difference-in-differences-like estimation [81] to address some concerns related to time-invariant unobserved confounders.

We define the base statistical model (eq. 1.1) by estimating the individual future GPA as a function of the individual average reciprocated cooperation (R_i, see methods section and Table 1.1) as:

$$GPA_{i1} = \beta_1 R_i + \beta_2 s_i + \beta_3 Rank_i + \beta_4 GPA_{i0} + \beta_5 G_i + \beta_6 A_i + \beta_7 TESC_i + \theta_c + e_i, \tag{1.1}$$

where, GPA_{i1} represents the GPA of the student i and e_i is the error term. The effect of reciprocated cooperation could differ between students with different levels of sent tokens and social status within their classrooms, besides both variables are arguable correlated with GPA (See SM3.1 and Table A.3). To control for these sources of variation, we include the average number of sent tokens (s_i) and the PageRank ($Rank_i$)–a network measure to proxy individual social ranking computed using the full cooperation network (See Method sections and Table 1.1). We also included traditional confounding variables, such as gender (G_i), percentage of class attendance (A_i), level of education of the guardian ($TESC_i = 1$ if the guardian completed secondary school, 0 otherwise), and fixed effects per classroom (θ_c).

Table 1.2 displays the estimation for the model defined by Equation 1.1, with a few variations in control variables. Column 1 shows the effect of reciprocated cooperation (R_i, presented as z-score in the whole sample for interpretability) within each classroom. We control for classrooms' average unobserved differences using fixed effects θ_i. We note that 5.4% of the explained variance (R^2 within) is due to the reciprocated cooperation within classrooms. However, to properly study the effect of reciprocated cooperation and to avoid omitted variable biases, we need to account for the individual average cooperation because, by definition, a higher average cooperation lead to a higher reciprocated cooperation (see Table 1.1 and Table A.3 Model 1). We also need to control for individual social status ($Rank_i$) measured as the PageRank network centrality [78; 82] (see Table 1.1 and Table A.3 Model 2) because a student with a higher social status will be the target of more cooperation leading to an increase in their reciprocated cooperation. Therefore, column 2 shows our model controlling for sent cooperation and social rank, and we observe that the three variables are significant and they explain 18.3% of the variation within classroom. Finally, column 3 shows our model controlling for the traditional confounding variables. We note that previous semester's GPA (Grades before measuring GPA_{i0}) quantifies individual previous accomplishment. Prior individual GPA provides a proxy for individual talent and controls by several time-invariant confounding variables that correlates with GPA, such as household income, practicing sports, among others. We also included attendance percentage, the education level of the student tutor, and students' sex. Thus, the total explained variance is 72.8% (See Supplementary SM 3.3 for a predicted v/s observed values plot for model 3 using both future GPA as dependent variables).

We find a positive and significant effect of reciprocated cooperation on GPA. More precisely, we find that an increase of one standard deviation in reciprocated cooperation is associated with an increase of 0.094 units in future GPA. We note that GPA in Chile goes from 1.00 to 7.00, and the average GPA of the first and the second analyzed semester are 5.87 ± 0.58 and 5.79 ± 0.57, respectively. Therefore, the average variation between both semesters is $\Delta_{GPA} = -0.080$. Thus, the reciprocated cooperation effect size (0.094) and the average decrease of GPA between the two periods (-0.080) are comparable. Indeed, the effect size is 117.5% of the average variation in GPA between both semesters.

However, some meaningful unobserved confounding variables could be affecting our results. For instance, changes in the classroom configuration, including new teachers, new students, or increasing/decreasing the school income. Also, changes at the individual level, such as a new job for the student's guardian or a house move, would impact household income and social capital outside of the classroom, respectively. Besides, other individual changes, such as a student's illness, a long trip, failing a school year, or family issues, would impact school engagement, among other outcomes. However, we assume that most of these unobserved variables are time-invariant within our study period spanning July to December 2017 (see Supplementary Notes 1).

Finally, to overcome all of the potential issues related to time-invariant unobserved confounders and provide evidence on the magnitude of the relationship between reciprocated cooperation and

	Dependent variable:		
	GPA (after measuring)		
	(1)	(2)	(3)
Reciprocated cooperation (z-score)	0.161***	0.394***	0.094***
	(0.03)	(0.05)	(0.03)
Sent cooperation (z-score)		-0.285***	-0.063**
		(0.04)	(0.03)
Rank (z-score)		0.145***	0.065***
		(0.04)	(0.02)
Grades (before measuring)			0.654***
			(0.02)
Attendance (%)			0.010***
			(0.00)
Tutor comp. sec. school (yes)			0.019
			(0.02)
Sex (Male)			-0.055**
			(0.03)
Fixed effects	Class-group	Class-group	Class-group
Observations	769	769	769
R^2	0.268	0.368	0.746
Adjusted R^2	0.222	0.327	0.728
R^2 within	0.054	0.183	0.672
F Statistics	41.079	54.010	209.438
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 1.2: OLS regressions for students' GPA after measuring cooperation. Note that these models show results for 769 students because we do not have the data for Tutor complete secondary school for 86 students.

GPA improvement, we use a treatment intensity difference-in-differences-like framework [81]. Here, our treatment intensity variable is individual average reciprocated cooperation (R_i), a continuous variable that induces variation at the individual level. Our specification is the following:

$$GPA_{it} = \beta_1 + \beta_2 T + \beta_3 R_i + \beta_4 A_{it} + \delta T \times R_i + \varepsilon_i + e_{it}, \tag{1.2}$$

where ε_i represents individual-level fixed effects (we note that individual-level fixed effects absorb classroom level fixed effects) and e_{it} is the error term. GPA_{it} is the GPA of student i in period t,

	Dependent variable:			
	GPA (Diff-in-Diff)			
	(1)	(2)	(3)	(4)
Reciprocity * Time	0.039***	0.038***		
	(0.01)	(0.01)		
Time	-0.081***	-0.094***	-0.102***	-0.114***
	(0.01)	(0.01)	(0.01)	(0.01)
Attendance (%)		0.008***		0.008***
		(0.00)		(0.00)
Top 20% of Reciprocity * Time			0.106***	0.100***
			(0.03)	(0.03)
Fixed effects	Individual	Individual	Individual	Individual
Observations	1710	1710	1710	1710
Students	855	855	855	855
R-squared	0.894	0.897	0.894	0.897
Adjusted R-squared	0.787	0.793	0.787	0.793
R-squared within	0.054	0.082	0.056	0.083
F Statistics	24.535	25.217	25.504	25.604
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1.3: Difference-in-differences-like estimation of GPA. Models 1 and 2 differ in the control variable "class attendance". In Models 3 and 4, we built a dummy variable for reciprocated cooperation. It takes value 1 if the reciprocated cooperation level for individual i is in the top 20%, and it takes value 0 otherwise.

T represents the semester and it takes values 0 (before measuring, $t = 0$) and 1 (after measuring, $t = 1$), R_i is the treatment intensity (reciprocated cooperation), A_{it} is the class attendance for both time periods, and finally, the diff-in-diff-like estimator is represented by δ. Table 1.3 shows four variations of the specification showed in eq. 1.2.

Therefore, we argue that our diff-in-diff-like estimator is consistent and robust with models in Table 1.2. Particularly interesting, is model 2 on Table 1.3. The diff-in-diff-like estimator is positive (0.038) and significant (p-value$<$ 0.01), above and beyond the effects of the unobserved time-invariant individual characteristics (individual fixed effects). The size effect of our diff-in-diff-like estimation (Model 2 Table 1.3) is nearly the 40% of the associative effect presented in Table

1.2 model 3.

In model 4 Table 1.3, we set a dummy variable for the reciprocated cooperation as 1 for all individuals in the top 20% of reciprocated cooperation and 0 otherwise. We observe a higher size effect of reciprocated reciprocity (0.100), suggesting that the effect is heterogeneous in the reciprocated reciprocity range. Indeed, as a robustness check, we estimate the effect of reciprocity on GPA by setting nine dummies for different thresholds of reciprocity, ranging from the bottom-10% to the top-10% of reciprocated cooperation (Table A.4). Despite that the average effect is of reciprocity is 0.038 (Table 1.3 Model 2), we found that the effect of reciprocity increases from the median to the top 10% reciprocity, confirming that the effect is more substantial for students over the top-20% of reciprocated cooperation.

1.4 Discussion

Individuals' position in their social networks is associated with educational outcomes at all ages [1–13]. Moreover, the literature on social learning states a positive relationship between a student's social network position and academic performance. There are three possible and non-exclusive explanations in this context: i) Central students get better GPAs through positive learning externalities from their social connections; ii) Higher GPA leads to a higher status, resulting in more central students in their social networks; iii) More talented students are more strategic in their interactions and learn faster, leading to increases in GPA that expand the gap with less talented students. Here, we provide evidence for the first explanation by studying cooperative patterns in elementary school students using a video game based on game theory and a diff-in-diff-like identification strategy.

Studying educational outcomes in elementary school is challenging because of the potential biases in survey measurement instruments. Therefore, we implemented a lab in the field approach to map students' social relationships in their classrooms. Students played a non-anonymous social dilemma on interconnected tablet computers through a friendly user interface where they had to choose how many tokens to share with their peers (Fig. 1-1 A). Thus, we mapped the entire student's cooperative network in their classroom (Fig. 1-1 B). We found that students mainly engage in three types of cooperative relationship: fully cooperative(Fig. 1-2, I), non-cooperative (Fig. 1-2, II), and relationships in which the cooperation is asymmetric (Fig. 1-2, III).

Then, we define reciprocated cooperation (Table 1.1) as the minimum between sent and received cooperation and found it improves elementary school students' GPA. We provided evidence on the positive and significant effect of reciprocated cooperation and GPA used both linear models (Table 1.2) and a difference-in-difference-like identification strategy (Table 1.3) exploiting an endogenous treatment intensity variable [81]. In the former, we show that the effect of received, sent, and reciprocated tokens survives even after controlling by confounding variables, such as previous GPA, sex, class attendance, and the educational level of the student's guardian. We find evidence supporting reciprocated cooperation as a predictor of future GPA. In the latter, the central assumption is that both control and treatment groups follow parallel trends in the absence of our treatment intensity variable, the reciprocated cooperation. Changes in GPA follow the same trend in the absence of reciprocated cooperation for all students. Some limitations exist here. Any change within the time interval studied in a confounding variable that affects both GPA and reciprocated cooperation could hinder our results. For instance, a change in family income may impact reciprocity through the popularity of a student. Also, it may affect GPA through getting a private professor or changes in the access to the internet [23]. Also, practicing sports may impact reciprocated cooperation through popularity, and also it may affect GPA [22; 83] at getting healthy. Thus, any change between the first and second semester of the academic year in which the experiment was implemented in the level of a guardian's student income or the student started to practice a sport could have impacted reciprocated cooperation and GPA. These changes would make the parallel trend assumption for our difference in difference specification unfulfilled because we cannot control or measure any of these changes. It is always impossible to ensure that parallel trends assumption is wholly fulfilled. Yet, we provide robust and consistent evidence of a directional effect from reciprocated cooperation to GPA improving under the premise of parallel trends is fulfilled (see Supplementary Notes 1). On the other hand, by controlling for unobserved time-invariant individual characteristics, we already account for all features that arguably do not change from one semester to the next such as talent. Therefore, a different statistical setting is needed to explore the role of talent stated in the third explanation mentioned before.

We found that the three social network measures–Social Rank, Sent cooperation, and Reciprocated cooperation (see Table 1.1)–account for 18.3% of the variance within classrooms, where $5,4\%$ is given by Reciprocated reciprocity (Table 1.2). The explained variance remains the same even after controlling for all time-invariant confounders (Table 1.3). As a robustness check, we explored the heterogeneity of the effect of reciprocity on GPA (Table A.4) and we found that the size effect of reciprocity increases from 0.039 to 0.100 for students belonging to the top-20% of reciprocated cooperation.

We consider this work to contribute to our understanding of the link between social networks and learning outcomes and how novel methodologies such as experimental game theory can help us in this endeavor.

From a methodological perspective, our results open new avenues for the role of game-theoretical and network-based tools to leverage relational information in elementary education [53; 84] bypassing the common biases of survey instruments. We contribute to the external validity of game-theoretic experiments in school context, producing a measure of the individual's social capital embodied directly in the student's relationships and indirectly by the configuration of the network as a whole [84]. Using this approach, we are able to map a representative social network for a group of young people with a common history [50].

Finally, from a policy perspective, the natural question to ask is what kind of intervention might improve students' academic performance by optimizing the potential benefits of cooperation. Our results, together with the lack of GPA homophily (see Supplementary Notes SM 1.2, Fig. A-1 and Table A.1), suggest that encouraging social relationships through, for instance, interventions of the spatial arrangement of the class [85], fostering community-school partnership [86; 87], an instructional design that promotes the formation of social ties in the classroom [10; 13; 35; 36; 88–92], or any other intervention that aims to support cooperation in classrooms [76; 93–96] are potentially fruitful alternatives to explore. Thus, we open the possibility for intervention by promoting relationships within the classroom that might significantly affect the academic achievements of primary students.

Bibliography

[1] T. T. Baldwin, M. D. Bedell, and J. L. Johnson, "The social fabric of a team-based M.B.A. program: Network effects on student satisfaction and performance," *Academy of Management Journal*, 1997.

[2] G. V. Caprara, C. Barbaranelli, C. Pastorelli, A. Bandura, and P. G. Zimbardo, "Prosocial foundations of children's academic achievement," *Psychological Science*, 2000.

[3] J. Bruun and E. Brewe, "Talking and learning physics: Predicting future grades from network measures and Force Concept Inventory pretest scores," *Physical Review Special Topics - Physics Education Research*, 2013.

[4] D. Gašević, A. Zouaq, and R. Janzen, "Choose Your Classmates, Your GPA Is at Stake!: The Association of Cross-Class Social Ties and Academic Performance," *American Behavioral Scientist*, 2013.

[5] D. Blansky, C. Kavanaugh, C. Boothroyd, B. Benson, J. Gallagher, J. Endress, and H. Sayama, "Spread of Academic Success in a High School Social Network," *PLoS ONE*, 2013.

[6] V. Ivaniushina and D. Alexandrov, "Anti-school attitudes, school culture and friendship networks," *British Journal of Sociology of Education*, vol. 39, pp. 698–716, jul 2018.

[7] C. Stadtfeld, A. Vörös, T. Elmer, Z. Boda, and I. J. Raabe, "Integration in emerging social networks explains academic failure and success.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, pp. 792–797, jan 2019.

[8] V. Kassarnig, E. Mones, A. Bjerre-Nielsen, P. Sapiezynski, D. Dreyer Lassen, and S. Lehmann, “Academic performance and behavioral patterns,” *EPJ Data Science*, vol. 7, p. 10, dec 2018.

[9] M. Berthelon, E. Bettinger, D. I. Kruger, and A. Montecinos-Pearce, “The structure of peers: The impact of peer networks on academic achievement,” *Res. High. Educ.*, vol. 60, pp. 931–959, Nov. 2019.

[10] J. Pulgar, C. Candia, and P. Leonardi, “Social networks and academic performance in physics: Undergraduate cooperation enhances ill-structured problem elaboration and inhibits well-structured problem solving,” *Physical Review Physics Education Research*, vol. 16, no. 1, p. 010137, 2020.

[11] C. Candia, A. Maldonado-Trapp, K. Lobos, F. Peña, and C. Bruna, “Disadvantaged students increase their academic performance through collective intelligence exposure in emergency remote learning due to covid 19,” *arXiv preprint arXiv:2203.05621*, 2022.

[12] C. Candia, J. Pulgar, and F. Pinheiro, “Interconnectedness in education systems,” *arXiv preprint arXiv:2203.05624*, 2022.

[13] J. Pulgar, D. Ramírez, A. Umanzor, C. Candia, and I. Sánchez, “Long-term collaboration with strong friendship ties improves academic performance in remote and hybrid teaching modalities in high school physics,” *arXiv preprint arXiv:2203.05638*, 2022.

[14] D. Halpern, *Social capital.* Polity, 2005.

[15] A. Claessens and M. Engel, “How Important Is Where You Start ? Early Mathematics Knowledge and Later School Success,” *Teachers College Record*, 2013.

[16] W. T. Dickens, I. Sawhill, and J. Tebbs, “The effecgs of investing in ealry education on economic growth,” tech. rep., The Brookings Institution, Washington, DC, 2006.

[17] D. Restuccia and C. Urrutia, “Intergenerational persistence of earnings: The role of early and college education,” *American Economic Review*, 2004.

[18] K. Ahmad and B. Naeema, “Influence of socioeconomic and educational background of parents on their children’ education in nigeria,” *International journal of scientific and research publications*, vol. 3, 2013.

[19] R. A. Marcon, “Positive relationships between parent school involvement and public school inner-city preschoolers’ development and academic performance,” *School Psychology Review*, 1999.

[20] T. H. Macan, C. Shahani, R. L. Dipboye, and A. P. Phillips, “College students’ time management: Correlations with academic performance and stress.,” *Journal of Educational Psychology*, vol. 82, no. 4, pp. 760–768, 1990.

[21] G. Curcio, M. Ferrara, and L. De Gennaro, “Sleep loss, learning capacity and academic performance,” 2006.

[22] A. Singh, L. Uijtdewilligen, J. W. R. Twisk, W. van Mechelen, and M. J. M. Chinapaw, “Physical activity and performance at school: a systematic review of the literature including a methodological quality assessment,” *Archives of pediatrics & adolescent medicine*, 2012.

[23] L. A. Jackson, A. von Eye, F. A. Biocca, G. Barbatsis, Y. Zhao, and H. E. Fitzgerald, “Does home internet use influence the academic performance of low-income children?,” *Developmental Psychology*, vol. 42, no. 3, pp. 429–435, 2006.

[24] J. Henrich, *The Secret of Our Success How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press, 2015.

[25] F. J. Gil-White and J. Henrich, “The Evolution of Prestige,” *Evolution and Human Behavior*, 2001.

[26] A. Pentland, *Social Physics: How Social Networks Can Make Us Smarter*. Penguin Books, 2015.

[27] D. W. Johnson and R. T. Johnson, *Learning together and alone: Cooperative, competitive, and individualistic learning.* Prentice-Hall, Inc, 1987.

[28] T. Roger and D. W. Johnson, "An overview of cooperative learning," *Creativity and collaborative learning*, pp. 1–21, 1994.

[29] B. Sacerdote, *Peer Effects in Education: How might they work, how big are they and how much do we know Thus Far?* 2011.

[30] P. Davies, *Paying for Education: Debating the Price of Progress.* 2018.

[31] S. Biancani and D. A. McFarland, "Social Networks Research in Higher Education," in *Higher Education: Handbook of Theory and Research*, ch. 4, pp. 151—-215, London: Springer, 2013.

[32] A. Calvó-Armengol, E. Patacchini, and Y. Zenou, "Peer effects and social networks in education," *Review of Economic Studies*, 2009.

[33] G. Csibra and G. Gergely, "Natural pedagogy as evolutionary adaptation," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 366, no. 1567, pp. 1149–1157, 2011.

[34] J. Henrich and F. J. Gil-White, "The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission," *Evolution and human behavior*, vol. 22, no. 3, pp. 165–196, 2001.

[35] D. A. McFarland, J. Moody, D. Diehl, J. A. Smith, and R. J. Thomas, "Network Ecology and Adolescent Social Structure," *American Sociological Review*, 2014.

[36] C. A. Rohrbeck, M. D. Ginsburg-Block, J. W. Fantuzzo, and T. R. Miller, "Peer-assisted learning interventions with elementary school students: A meta-analytic review," 2003.

[37] P. Kollock, "Social dilemmas: The anatomy of cooperation," *Annual review of sociology*, vol. 24, no. 1, pp. 183–214, 1998.

[38] J. L. Moreno, *Who shall survive?: A new approach to the problem of human interrelations.* Washington: Nervous and Mental Disease Publishing Co, aug 1934.

[39] J. D. Coie, K. A. Dodge, and J. B. Kupersmidt, *Peer group behavior and social status.* Cambridge University Press, 1990.

[40] T. Mouw, "Estimating the Causal Effect of Social Capital: A Review of Recent Research," *Annual Review of Sociology*, 2006.

[41] J. W. Neal, "Why Social Networks Matter: A Structural Approach to the Study of Relational Aggression in Middle Childhood and Adolescence," *Child & Youth Care Forum*, 2007.

[42] M. P. McCormick and E. Cappella, "Conceptualizing academic norms in middle school: A social network perspective," *The Journal of Early Adolescence*, 2015.

[43] E. D. de Leeuw, "Improving data quality when surveying children and adolescents: Cognitive and social development and its role in questionnaire construction and pretesting," in *Report prepared for the Annual Meeting of the Academy of Finland: Research programs public health challenges and health and welfare of children and young people*, pp. 10–12, 2011.

[44] T. F. Van de Mortel *et al.*, "Faking it: social desirability response bias in self-report research," *Australian Journal of Advanced Nursing, The*, vol. 25, no. 4, p. 40, 2008.

[45] N. Borgers, E. de Leeuw, and J. Hox, "Children as respondents in survey research: Cognitive development and response quality 1," *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, vol. 66, pp. 60–75, Apr. 2000.

[46] E. N. Banister and G. J. Booth, "Exploring innovative methodologies for child-centric consumer research," *Qualitative Market Research: An International Journal*, vol. 8, pp. 157–175, Jan. 2005.

[47] J. Barker and S. Weller, ""is it fun?" developing children centred research methods," *Int. J. Sociol. Soc. Policy*, vol. 23, pp. 33–58, Jan. 2003.

[48] M. Fleer and A. Ridgway, eds., *Visual Methodologies and Digital Tools for Researching with Young Children: Transforming Visuality*. Springer, Cham, 2014.

[49] K. Kyritsi, "Doing research with children: Making choices on ethics and methodology that encourage children's participation," *J. Child. Stud.*, pp. 39–50, July 2019.

[50] Z. Wang, M. Jusup, R. W. Wang, L. Shi, Y. Iwasa, Y. Moreno, and J. Kurths, "Onymity promotes cooperation in social dilemma experiments," *Science Advances*, 2017.

[51] C. F. Camerer and R. M. Hogarth, "The effects of financial incentives in experiments: A review and capital-labor-production framework," *Journal of risk and uncertainty*, vol. 19, no. 1-3, pp. 7–42, 1999.

[52] B. Wellman and K. A. Frank, "Network capital in a multilevel world: Getting support from personal communities," in *Social capital*, pp. 233–273, Routledge, 2017.

[53] R. S. Burt, "The Network Structure Of Social Capital," *Research in Organizational Behavior*, 2000.

[54] M. S. Granovetter, "The Strength of Weak Ties," *American Journal of Sociology*, 1973.

[55] P. V. Marsden and K. E. Campbell, "Measuring Tie Strength," *Social Forces*, vol. 63, p. 482, dec 1984.

[56] S. P. Borgatti and M. G. Everett, "A graph-theoretic perspective on centrality," *Social networks*, vol. 28, no. 4, pp. 466–484, 2006.

[57] A. Clauset, M. E. Newman, and C. Moore, "Finding community structure in very large networks," *Physical review E*, vol. 70, no. 6, p. 066111, 2004.

[58] J. Berg, J. Dickhaut, and K. McCabe, "Trust, reciprocity, and social history," *Games and economic behavior*, vol. 10, no. 1, pp. 122–142, 1995.

[59] E. Fehr and S. Gächter, "Cooperation and punishment in public goods experiments," *American Economic Review*, vol. 90, no. 4, pp. 980–994, 2000.

[60] E. Ostrom, R. Gardner, J. Walker, J. M. Walker, and J. Walker, *Rules, games, and common-pool resources.* University of Michigan Press, 1994.

[61] C. Rodriguez-Sickert, "Homo economicus," in *Handbook of economics and ethics*, Edward Elgar Publishing, 2009.

[62] C. F. Camerer and E. Fehr, "Measuring social norms and preferences using experimental games: A guide for social scientists," *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*, vol. 97, pp. 55–95, 2004.

[63] J. Henrich, S. J. Heine, and A. Norenzayan, "Most people are not weird," *Nature*, vol. 466, no. 7302, pp. 29–29, 2010.

[64] J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, *et al.*, ""economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies," *Behavioral and brain sciences*, vol. 28, no. 6, pp. 795–815, 2005.

[65] J.-C. Cárdenas and E. Ostrom, "What do people bring into the game? experiments in the field about cooperation in the commons," *Agricultural systems*, vol. 82, no. 3, pp. 307–326, 2004.

[66] C. Rodriguez-Sickert, R. A. Guzmán, and J. C. Cárdenas, "Institutions influence preferences: Evidence from a common pool resource experiment," *Journal of Economic Behavior & Organization*, vol. 67, no. 1, pp. 215–227, 2008.

[67] C. F. Camerer, *Behavioral game theory: Experiments in strategic interaction.* Princeton university press, 2011.

[68] J. Carpenter and E. Seki, "Do social preferences increase productivity? Field experimental evidence from fishermen in Toyama Bay," *Economic Inquiry*, 2011.

[69] Y. Algan, Y. Benkler, M. F. Morell, and J. Hergueux, "Cooperation in a peer production economy - experimental evidence from Wikipedia," *Working Paper*, 2013.

[70] E. Fehr and A. Leibbrandt, "A field study on cooperativeness and impatience in the Tragedy of the Commons," *Journal of Public Economics*, 2011.

[71] S. Gelcich, R. Guzman, C. Rodríguez-Sickert, J. C. Castilla, and J. C. Cárdenas, "Exploring external validity of common pool resource experiments: Insights from artisanal benthic fisheries in Chile," *Ecology and Society*, 2013.

[72] A. Hopfensitz and J. Miquel-Florensa, "Mill ownership and farmer's cooperative behavior: The case of Costa Rica coffee farmers," 2017.

[73] M. I. Rivera-Hechem, R. A. Guzmán, C. Rodríguez-Sickert, and S. Gelcich, "Effects of experience with access regimes on stewardship behaviors of small-scale fishers," *Conservation Biology*, vol. 35, no. 6, pp. 1913–1922, 2021.

[74] S. D. Levitt and J. A. List, "What do laboratory experiments measuring social preferences reveal about the real world?," *Journal of Economic perspectives*, vol. 21, no. 2, pp. 153–174, 2007.

[75] S. D. Levitt and J. A. List, "On the generalizability of lab behaviour to the field," *Canadian Journal of Economics/Revue canadienne d'économique*, vol. 40, no. 2, pp. 347–370, 2007.

[76] M. A. Nowak, "Five rules for the evolution of cooperation," *science*, vol. 314, no. 5805, pp. 1560–1563, 2006.

[77] J. K. Goeree, A. Riedl, and A. Ule, "In search of stars: Network formation among heterogeneous agents," *Games and Economic Behavior*, vol. 67, no. 2, pp. 445–466, 2009.

[78] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web.," tech. rep., Stanford InfoLab, 1999.

[79] D. S. Karlan, “Using experimental economics to measure social capital and predict financial decisions,” 2005.

[80] J. D. Angrist and J.-S. Pischke, *Mostly harmless econometrics*. Princeton university press, 2008.

[81] D. Acemoglu, D. H. Autor, and D. Lyle, “Women, war, and wages: The effect of female labor supply on the wage structure at midcentury,” *Journal of political Economy*, vol. 112, no. 3, pp. 497–551, 2004.

[82] E. E. Bruch and M. Newman, “Aspirational pursuit of mates in online dating markets,” *Science Advances*, vol. 4, no. 8, p. eaap9815, 2018.

[83] D. I. Rees and J. J. Sabia, “Sports participation and academic performance: Evidence from the national longitudinal study of adolescent health,” *Economics of Education Review*, vol. 29, no. 5, pp. 751–759, 2010.

[84] C. M. Lakon, D. C. Godette, and J. R. Hipp, “Network-Based Approaches for Measuring Social Capital,” in *Social Capital and Health*, pp. 63–81, New York, NY: Springer New York, 2008.

[85] S. J. Salend, “Facilitating friendships among diverse students,” *Intervention in School and Clinic*, 1999.

[86] S. J. Haines, J. M. Gross, M. Blue-Banning, G. L. Francis, and A. P. Turnbull, “Fostering family–school and community–school partnerships in inclusive schools: Using practice as a guide,” *Research and Practice for Persons with Severe Disabilities*, vol. 40, no. 3, pp. 227–239, 2015.

[87] N. Ratcliff and G. Hunt, “Building teacher-family partnerships: The role of teacher preparation programs.,” *Education*, vol. 129, no. 3, 2009.

[88] R. E. Slavin, "Cooperative learning in elementary schools," *Education 3-13*, vol. 43, no. 1, pp. 5–14, 2015.

[89] R. S. Newman, "How self-regulated learners cope with academic difficulty: The role of adaptive help seeking," *Theory into practice*, vol. 41, no. 2, pp. 132–138, 2002.

[90] R. E. Slavin, E. A. Hurley, and A. Chamberlain, "Cooperative learning and achievement: Theory and research," *Handbook of psychology*, pp. 177–198, 2003.

[91] R. E. Slavin, "Instruction based on cooperative learning," in *Handbook of research on learning and instruction*, pp. 358–374, Routledge, 2011.

[92] J. Pulgar, C. Rios, and C. Candia, "Physics problems and instructional strategies for developing social networks in university classrooms," *arXiv preprint arXiv:1904.02840*, 2019.

[93] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *science*, vol. 211, no. 4489, pp. 1390–1396, 1981.

[94] R. L. Trivers, "The evolution of reciprocal altruism," *The Quarterly review of biology*, vol. 46, no. 1, pp. 35–57, 1971.

[95] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, no. 7063, p. 1291, 2005.

[96] K. Sigmund, C. Hauert, and M. A. Nowak, "Reward and punishment," *Proceedings of the National Academy of Sciences*, vol. 98, no. 19, pp. 10757–10762, 2001.

[97] I. Smirnov and S. Thurner, "Formation of homophily in academic performance: Students change their friends rather than performance," *PLOS ONE*, vol. 12, p. e0183473, aug 2017.

[98] C. Candia, S. Encarnação, and F. L. Pinheiro, "The higher education space: connecting degree programs from individuals' choices," *EPJ Data Science*, vol. 8, no. 1, p. 39, 2019.

Chapter 2

Cooperative relationships and bullying/victimization in the classroom: insights from an experimental approach

Abstract

Positive effects of cooperative learning on students have been found and replicated in many different cases and contexts, from general outcomes like greater effort to achieve, better relationships between them, and improved psychological health to more specific ones like lowering bullying and victimization in the classroom. However, the general relationship between bullying and cooperation, and particularly how bullying-involved students have different cooperative dispositions in a classroom small scale cooperative network is still widely unexplored. Using a video-game-like interface, we implement a novel game-theoretical dyadic non-anonymous social dilemma in 47 public primary school classrooms in Chile to map the underlying cooperative network for each class. We also collect data from school records and a self-report instrument on bullying behavior. We characterize how bullies, victims, and bully/victims differ in their cooperative behavior. Using multilevel modeling and node-level topological measures of the cooperative network, we consistently found that victims and bully/victims are negatively related to receiving cooperation. This relation holds when we control for demographic characteristics and contextual effects over the classroom. These results open new applications and expansions of both methods and findings to understand and design better measurements and interventions of cooperative behavior and students involved in bullying.

2.1 Introduction

Cooperation has been at the heart of our success as a species since the first humans lived in small bands of hunter-gatherers [1–4]. In modern societies, cooperation in small-scale groups–analogous to upper paleolithic bands–places a crucial role in the success of such groups and their members. Some remarkable cases are: sustainability of common-pool resources exploited by small communities [5], use of collective intelligence to effectively tackle group tasks within organizations [6], and cooperative learning in schools [7].

Scholars in the cooperative learning community have focused on theories, interventions (i.e., change procedures and systems of incentives in the school and its actors), and impact analysis. They have consistently found and replicated a positive effect between a cooperative learning environment in the classroom and learning and academic achievements at different levels (e.g., [8]). They have caught other novel and promising effects of cooperative learning implementation in classrooms like a decrease in bullying and victimization (e.g., [9]). However, particular mechanisms and relational and individual-level understanding of how bullying and a cooperative learning environment, and more generally, cooperative behaviors and dispositions between students and cooperative structure in the classroom relate, are primarily unknown.

Bullying Behavior

Bullying behavior is defined when a person is a victim of constantly aggressive behavior from one or more peers over time and without the chance to defend himself, implying a power imbalance [10; 11]. This seminal work focused on defining profiles for victims and perpetrators of this behavior [10–12], and differentiating children who were both involved in bullying situations as victims and perpetrators, calling them "provocative victims." More research consistently confirmed this third type of profile directly involved in bullying, youths who are identified as bullies and as victims simultaneously, usually referred to as bully/victims [13–16].

However, by its own previous definition, bullying behavior is aggression between peers that operates within relationships of different kinds inside a group[17]. Thus, not only individual characteristics such as aggressive dispositions but also social relationships become crucial to a better understanding of it. Social network methods have become more prevalent in the study of the bullying phenomenon [17; 18] in order to capture this relational dimension. Furthermore, several authors highlight that bullying is a "group process" [19]. Two important implications derive from this idea: Peers that are not directly involved (usually called bystanders or non-involved) can also play a crucial role, whether assisting or reinforcing bullies, walking out of the situation, or stepping up to defend the victim. Also, group features like social climate and learning environments can affect bullying.

Previous research indicates that when comparing the peer relations of bullies, victims, and bully-victims to the rest of non-involved children, victims tended to have fewer social skills and were more likely to be socially marginalized (i.e., isolated or have few playmates), bully-victims were more aggressive and more likely to have fewer playmates, and bullies were more aggressive and had more leadership skills and larger social clusters than non-involved children [20].

Bullying and Cooperation

Just a handful of studies have explored relations between bullying and cooperation. For instance, in a surveyed social dilemma run in a private residential school among middle school students, the intervention of a bystander in a bullying situation was not efficient and even dangerous for stopping bullying [21]. Group intervention could be more effective, but it was unlikely that others subjects would want to cooperate. Then, the individual best response strategy (and therefore theoretical Nash Equilibrium) was not to defend the victim, even though everyone would be better if there were a collective action to stop bullying [21]. Another example is the work of Schuster [22], who runs, in a small sample of students, a prisoner's dilemma where she sets a condition of changing hypothetical targets (receivers) to being victimized students and taking into account bullying conditions of players too.

Cooperation, cooperative relations, and Experimental Game Theory

Indeed, not all group members benefit from cooperation in the same way, and cooperative dispositions between members vary by many factors. Cooperative network structures emerge and change, and positions in that network are particularly important because, for instance, allow students to reap the benefits of social learning and improve academic performance[23; 24].

In order to capture this ecological complexity of cooperation between classmates, we posit a novel method to elicit the underlying "cooperative network" using an experimental approach based on a game-theoretical paradigm. We use experimental game theory based on the premise that cooperative social norms that are prevalent in the real world can penetrate laboratory (and lab-in-the-field) behavior [25]. Concretely, there are previous studies that found that individuals who are more cooperative in the real world also behave more cooperatively in the lab [26; 27]. Also, groups who achieve higher levels of cooperation in the real world attain higher levels of aggregate cooperation when playing a social dilemma in the lab [26; 28–31].

The experimental literature has slowly moved to study real-world populations (e.g., [32]) and involve games implemented in the field which stylistically represent social dilemmas that agents face in real-life (e.g, [33; 34]). Other social science disciplines have also proposed using games that depart from classical protocols derived from experimental economics. Therefore, taking advantage of revealing individuals' preferences in ways that observational and survey or interview data cannot. At the same time, thoughtfully altering design and framing in a way where outcomes do provide insights into real-world behavior in that particular situation and context [35].

In our study, we focus specifically on the nature of cooperative relationships of students within a given class. For this purpose, we design a game where students face a dyadic social dilemma [36; 37] to gather information about students' dispositions to cooperate with each of their classmates.

A social dilemma is a scenario in which individual and collective interests collide. There are incentives to maximize individuals' payoffs, while this can only be achieved by generating a sub-optimal collective performance [38].

In our specification of the dyadic social dilemma, once two class members are matched, they simultaneously decide how many tokens they will send to their partners out of an endowment of ten tokens. Each token a student sends to his/her partner is multiplied by a factor of two. The total payoff for a student in each social exchange is given by the tokens he receives from his partner, aggregated with the tokens he/she keeps for him/herself. Although both players could gain from this exchange: for instance, if both players send all of their tokens to their partner, both will end up with twenty tokens, doubling their initial allocation and both being in a better situation than both keeping all their tokens. However, in any case, the unilateral best response (to maximize payoff, meaning getting more tokens) for any player is never to send tokens. Then, Nash Equilibrium of the game is that both players send zero tokens.

Students played synchronously in dyads using connected tablets with a simple but engaging video-game-like design. All possible pair combinations within the classroom have to play the game once, i.e., a round-robin design. Then, a complete weighted cooperative network can be elicited by using each decision of sent tokens from student A to student B as a weighted directed edge between Student A and Student B. The use of tablets and a friendly interface might help to mitigate cognitive barriers [39] and induce greater levels of engagement. In addition, the use of a behavioral instrument -versus a peer nomination survey- is also particularly appropriate given the developmental stage of the subjects in our sample.

To characterize the nature of each student's cooperative relationships, we compute two straightforward node-level standard topological measures of network centrality, weighted Instrength and weighted Outstrength. Instrength can be interpreted as the average disposition to cooperate from their peers towards him/her by receiving tokens from them, and Outstrength as the student's average disposition to cooperate with the rest of their peers by sending tokens. We also compute a dyadic measure of reciprocity posited by Squartini et al. [40], Reciprocity strength. Reciprocity strength measures, on average, how many tokens that the agent sends within each game are reciprocated by their classmates, i.e., a student with high reciprocity strength is a student who is involved in many relationships with high levels of mutual cooperation.

In order to identify the distinct nature of cooperative relationships for Bullies, Victims, and Bully/Victims, we explored how Instrength, Outstrength, and Reciprocity Strength vary across bullying categories. We use multilevel models to control for class-level effects. Our results show that the different types do not differ in terms of the average cooperative dispositions toward the rest of his/her classmates (out-strength), indicating that this variable by itself is not very informative. We also show that cooperation towards a student (Instrength) is (compared to non-involved) lower for victims and bully/victims but not for bullies. This negative association between being a victim or a bully-victim and in-strength is "robust" to other controls (GPA, attendance) and to controls at the classroom level, both demographic (course GPA, course attendance) and also to bullying contextual effects (Average bullying and victimization score of the class). Finally, we show that being categorized as a bully is negatively and significantly associated with reciprocity strength. This association is robust to bullying contextual effects but vanishes when controlled by demographic variables. Contextual effects are significant, with class level of victimization negatively associated with Instrength and Reciprocity strength, and class level of bullying positively associated with Reciprocity strength.

2.2 Methods

2.2.1 Participants (Sample)

This research took place in a public school district of a municipality in a central metropolitan area of Chile. We implemented our data collection over 47 classrooms from 14 schools in this district, from grades 3rd, 4th, and 5th, with an approximate average age of 9.78 years and a gender ratio of 57% girls and 43% boys. Also, data from student school records were collected. In Chile, elementary and middle schools are usually part of the same institution, and it is quite common that most of the students in the sample have been classmates since first grade.

2.2.2 Procedure

Data was collected during the 2017 school year as part of a larger study, implementing our game and questionnaires simultaneously to a whole class per time. More details can be found in the supplementary material. Active consent for participation of schools in this activity was obtained from principals and class teachers. Passive consent was obtained from parents/guardians. Our data collection methods and the experimental protocol were approved on May 4, 2016, by the human subjects review board of Universidad del Desarrollo (IT15I10079).

2.2.3 Instruments and Measures

Bullying and Victimization Score

Illinois Bullying Scale was used to measure bullying and victimization scores from students. Of the three constructs from the original instrument, we only used two, Bully Score and Victim Score. Confirmatory Factor Analysis and Item Reliability were checked (see more details in SM).

Self-Reported Bully Score (SRBS).

We used the University of Illinois Bully Scale (UIBS [41]) self-report survey to assess bullying behavior, which includes teasing, social exclusion, name-calling, and rumor spreading [42; 43]. Students were asked to indicate how often in the past 30 days they have engaged in each behavior (e.g., “I teased other students”). Response options were “Never”, “Almost never”, “Sometimes” and “Almost always”, and coded 1 through 4 accordingly. The construct validity of this scale has been supported via exploratory and confirmatory factor analysis [41]. In a different study with a similar population, Varela et al. [44] reported a mean of $M = 1.50$ and a $SD = 0.74$ as raw scores. In this study, the level of reliability achieved was over our defined Cronbach’s alpha acceptance cutoff ($\alpha = 0.828$).

Self-Reported Victimization Score (SRVS).

Victimization from peers was assessed using the University of Illinois Victimization Scale (UIVS; [41]). Students answer how often the following things have happened to them in the past 30 days (i.e., "Other students picked on me," and "I got hit and pushed by other students"). Response options were "Never," "Almost never" "Sometimes" and "Almost always". The scale was not significantly correlated with the Illinois Bullying scale ($r = .12$), providing evidence of discriminant validity [43]. This scale also converged with peer nomination data [43]. In another study with a similar population [44], mean score and SD of 1.92 (0.99) were reported. For this study, the level of reliability (Cronbach's alpha) achieved was $\alpha = 0.802$, which is over our acceptance level.

Bullying Categories

A combination of Self-Reported Victimization Score and Self-Reported Bullying Score were used to define four mutually exclusive Bullying-involvement subtypes. These are: Bully, Victim, Bully-Victim, and Non-Involved. Standardization by gender was performed for SRBS and SRVS of all students by subtracting its corresponding gender-specific mean of the whole sample and then dividing by gender-specific standard deviation of that item. A 0.5 standardized score (0.5 SD over whole sample gender mean) was used as a cutoff to identify Bullies and Victims, to be consistent with other studies that differentiate among these four subtypes [13–16]. The specific criteria were that if students had 0.5 SD above on the gender-standardized SRBS and 0.5 above on gender-standardized SRVS, they were classified as Bully-Victim. If they had above 0.5 SD on SRBS and less than 0.5 SD on SRVS, they were categorized as Bully. If they had above 0.5 SD on SRVS and less than 0.5 SD on SRBS, they were classified as victim, and if they had less than 0.5 on both scores were classified as Non-Involved. Standardization by gender was used because girls usually have lower average scores than boys in bullying scores and, in this case, victimization scores. So we want to induce some gender balance in categories. In this sample, mean SRBS score for boys was 1.90 and for girls was 1.55, and doing a simple t-test on difference in means shows a significant difference ($0.35, t = 10.06$), same for SRVS, mean for boys 2.49, for girls 2.28, difference of $0.21, t = 3.89$.

The Game

Students play a modified version of a prisoner dilemma game. Each player is paired with a different classmate in every round. Every dyad has a one-shot interaction, with $n-1$ rounds for a class of n students. Given our focus on the relational dimension, we depart from the standard protocol, so players are informed with whom they are paired in every round (a non-anonymous or onymous frame [45]), and they simultaneously play the "Allocator" and the "Receiver" role. We give a general aggregate result over the total sent and received tokens after the game ends as general feedback, not giving results on each round.

All players receive an endowment of ten tokens each round, and they have to decide if they want to send tokens to the Receiver, choosing from zero to ten, keeping the rest. For every sent token, their classmate receives two tokens. So, the total number of tokens that a student gets in a round equals the number of tokens they kept (in their allocator role) plus twice the number of tokens sent by the other player to her. Hence, the couple maximizes the total number of tokens earned when they both cooperate (mutually give all their tokens to each other, in total, 20 tokens per student), but a single student maximizes his tokens when he defects (keep all his (10) tokens) and the other player cooperates with all of his tokens with him (10 tokens doubled, 20 tokens, giving a total of 20+10=30 tokens to the first student and zero tokens for the second one).

Network Elicitation and Network measures

We use the game described in the previous section to map a cooperation network for each classroom. In each dyadic interaction, tokens sent from student i to student j are represented with a directed edge from i to j with weight equal to the number of sent tokens.

For the regression analysis, we build the following individual-level topological measures from the elicited network.

Weighted in-degree or Instrength, calculated as the sum of all weights of edges directed to a student, and equivalent to the total received cooperation by a student.

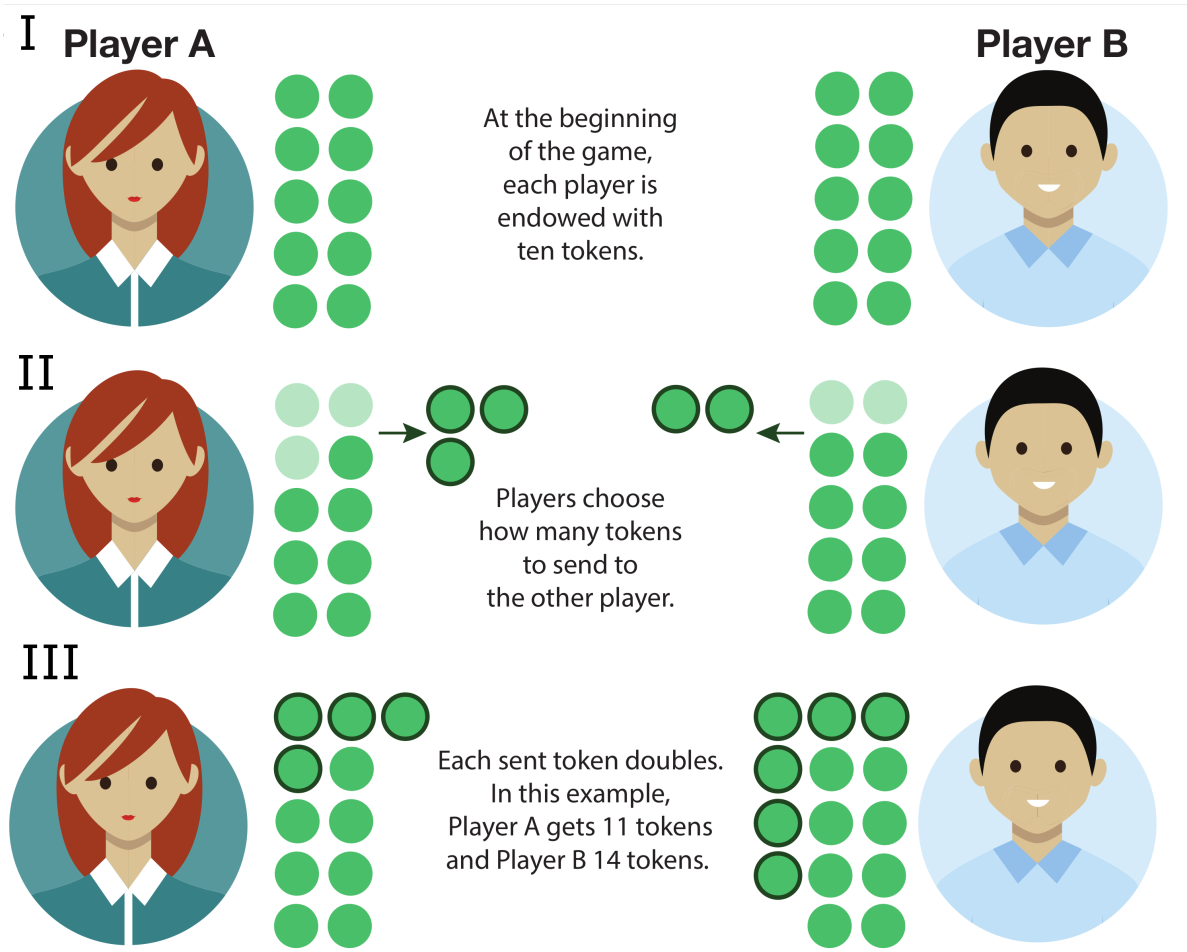

Figure 2-1: The Dyadic Social Dilemma. Example: I) Both players are endowed with 10 (ten) tokens. B) Simultaneously, Player A decides to send 3 tokens, and Player B decides to send 2 tokens. III) Sent tokens are doubled. Player A receives 4 tokens and second player receives 6 tokens. Note: Reprinted from [46]

Weighted out-degree or Outstrength, calculated as the sum of all weights of edges directed from a student to all other classmates, and equivalent to the total sent cooperation by a student.

Reciprocity Strength, calculated as the sum of all reciprocated weights of a student. Reciprocated weight is the minimum weight in a dyad.

In order to have comparable quantities between classrooms with different number of classmates, we divided each quantity by $n-1$ where n is the number of students in that class that played the game. Therefore, values for every student can only take values between 0 and 1.

Reciprocity was initially posited as a topological measure of a weighted network by Squartini et al. [40].

A

B

	K	L	M	N
K	0	0	5	10
L	0	0	2	5
M	10	3	0	10
N	10	10	0	0

C

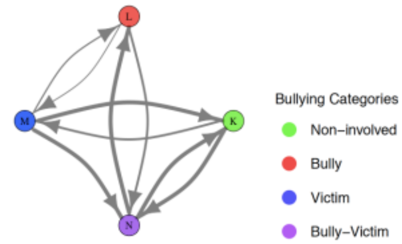

Figure 2-2: Example from a hypothetical class with $N = 4$ students. A) Tablet computer platform. B) Adjacency matrix where w_{ij} represents the tokens sent from student i to student j in the dyadic game. C) Cooperative weighted network that graphically represents the information displayed in the adjacency matrix. The width of the edge represents the amount of sent tokens and the node's size, the number of tokens each student ended up with at the end of the game. Node color indicates the bullying category assigned to the student.

Multilevel Methods

Educational data usually have a nested structure, so multilevel models are usually a suggested solution to deal with this non-independence and model the data structure. Also, multilevel modeling allows us to model variance even if we are controlling over a higher-level variable, which is not possible with, for example, a multivariate regression model with fixed effects. Given that students' behavior in the game was measured in the classroom toward other students from the same classroom without anonymity, our cooperative variables are naturally nested and non-independent between classrooms, and we are also interested in bullying and victimization effects at the classroom level, we are looking for potential contextual effects, so using multilevel modeling to analyze our data seems an appropriate choice. Furthermore, with multilevel models, it is possible to add higher-level variables to the model to estimate its effects. This is also possible via aggregating variables from a lower level, and then we can estimate "contextual effects".

For many reasons, multilevel models are usually centered and standardized. Most common centering practices are grand mean centering and cluster/group mean centering. We decided to standardize all our independent variables following [47] (using the R package arm), which is a grand mean centering for numeric variables, proportion centering for binary categorical variables, and standardization by dividing each variable over two standard deviations of that variable.

These models were implemented in R, using the lme4 package [48]. Other packages were used to clean and tide data (Tidyverse [49]), run confirmatory factor analysis (lavaan [50]) and to plot results and create tables and figures (sjPlot, [51]). Missing data were handled using listwise deletion, removing subjects from all the modeling. All individual student level variables were aggregated by classroom to form level 2 variables.

2.3 Results

2.3.1 Descriptive Results

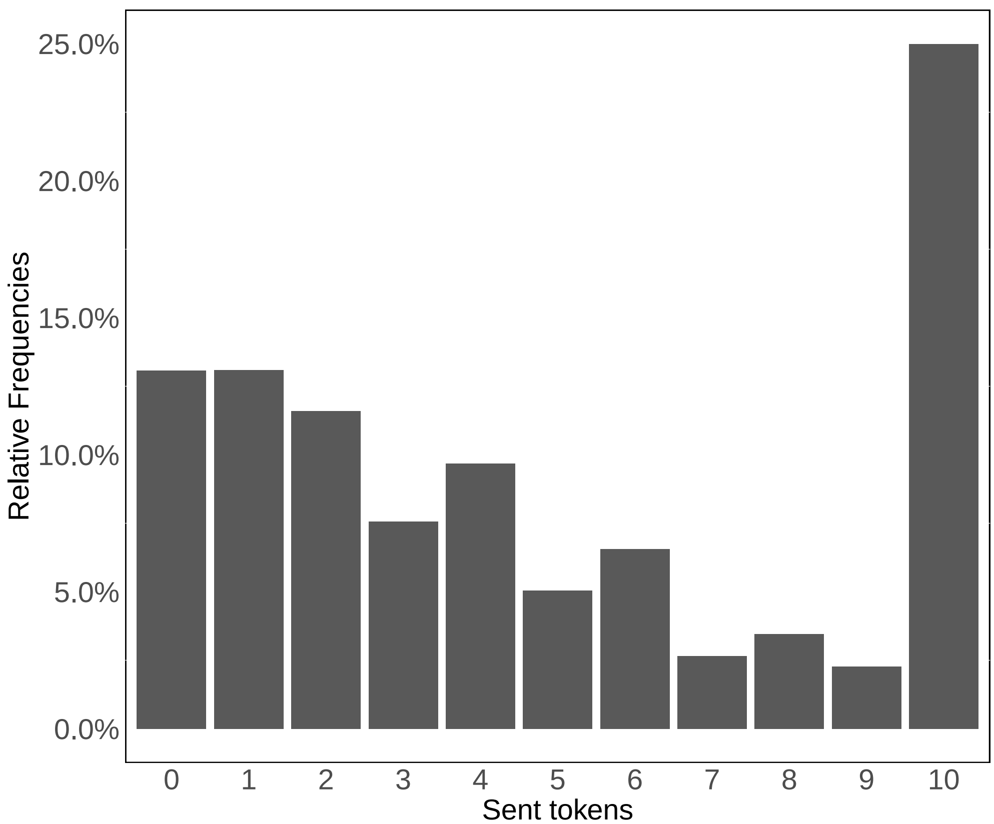

Figure 2-3: **Individual behavior in the dyadic social dilemma:** Histogram of sent tokens (individual level).

We first report general behavior towards the game. First, sent token distribution. Given that each classroom can have a different number of rounds, given by $(n-1)$, where n is the number of

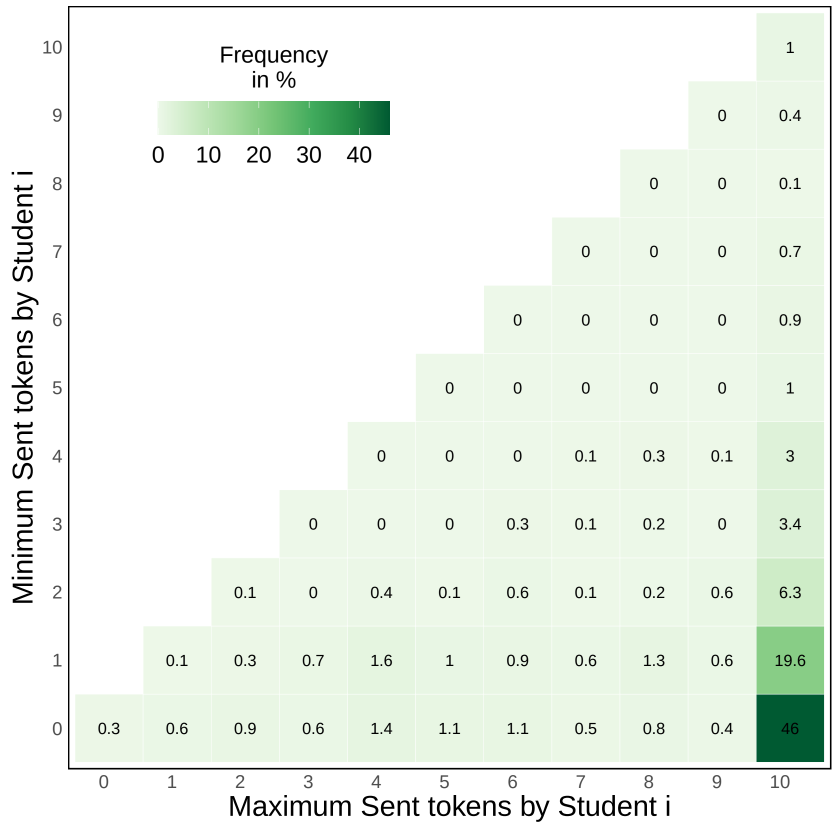

Figure 2-4: **Histogram of min-max sent tokens by students** .

kids that were authorized to participate in this activity that day. Then, figure 2-3 is weighted by the student, where regardless of the different number of rounds given by class size, each student contributes the same uniform weight to the model. In any case, the general result does not change much by taking into account this weight or not.

We plot a matrix-like heatmap where "hot spots" are dyads that are more frequent (towards red) and less frequent (towards yellow). This is a good visual queue, and it is fairly clear to the eye that more frequent dyads are mainly clustered in three places: Dyads with a lot of mutually highest cooperation (10-10 dyads, for example), dyads with mutually low cooperation (0-0 and 1-1, 1-0, 0-1, and even 0-2,1-2,2-2) and highly asymmetric dyads (10-0, 10-1, 10-2). The predicted Nash equilibrium from using game theory to "solve" the game, assuming rationality of preferences and individual payoff maximization measured only in tokens, would be sending zero tokens for both players. However, despite 0-0 having an important frequency (almost 4.5%), it is just half of full

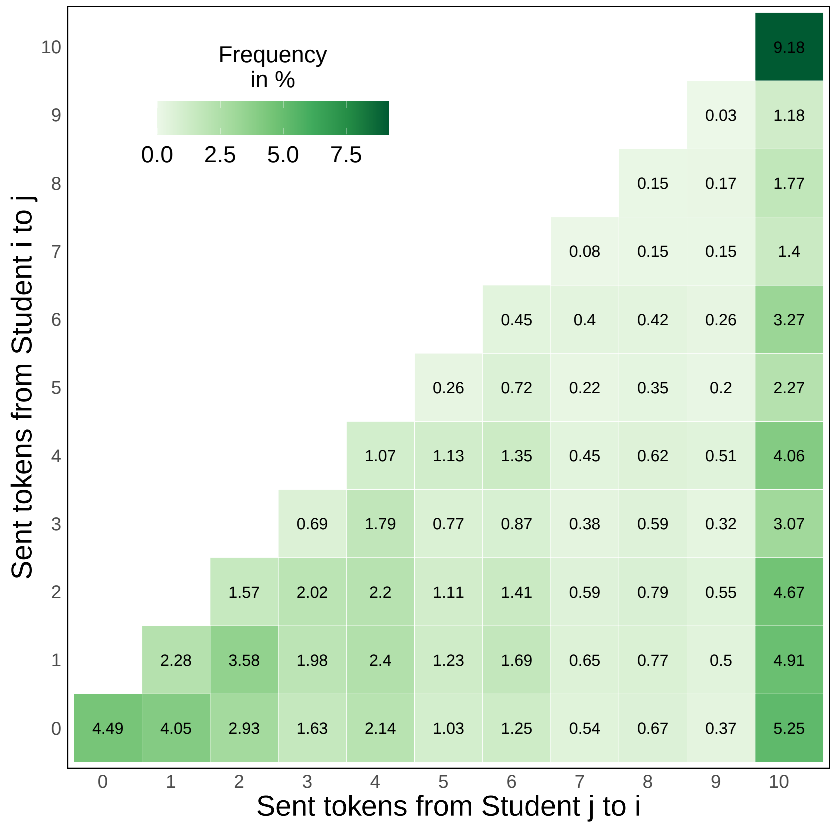

Figure 2-5: **Dyadic behavior in the social dilemma:** Histogram of experimental outcomes at the dyadic level over all sample.

mutually cooperative dyads (10-10) (9.18%). Then, this heatmap clearly shows that most of the plays do not fall into mutual non-cooperation.

As we explained in the Methods section, from the results of the implementation of the experiment using a set of networked tablets, we elicit the structure of cooperation for each class. Below we plot one example classroom network for an illustrative class:

2.3.2 Multilevel Analysis

To account for the amount of variance nested in student and classroom levels, we run null multilevel models on each of the four cooperative network metrics with random intercepts for Classrooms. Although it would have been interesting to include a separate level for schools, the low number of them made it not feasible to estimate reliable models with that specification. Distributions of these

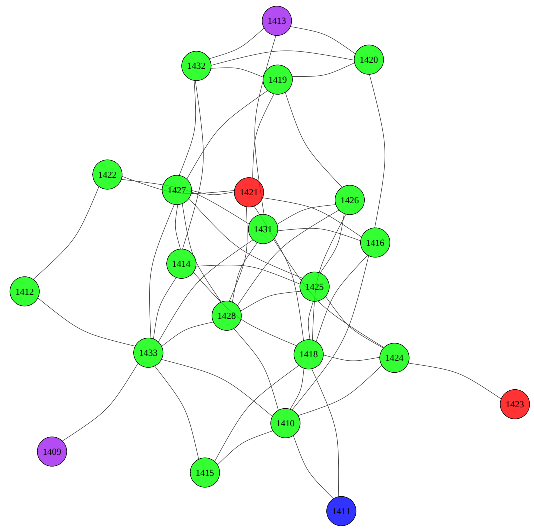

Figure 2-6: **Classroom reciprocity network:** Red nodes correspond to bullies, purple nodes to bully/victims, light blue to victims, and green nodes correspond to non-involved. Linked nodes correspond to relationships in which both members of the dyad sent more tokens to their partners than what they sent on average to the rest of the class members, i.e., they have a relation of mutual cooperation. In this case, victim and bully/victims are mostly located on the periphery of this network.

three variables (Instrength, Outstrength, and Reciprocity Strength) were quasi-normal, allowing us to run linear multilevel models. Intra-Class Correlation (ICC) is 0.454 for Instrength, 0.116 for Outstrength, and 0.331 for Reciprocity Strength in this first round of models. These variance partitions verified that a multilevel analysis is pertinent.

Then, we explored whether self-perceived bullying-involvement categories had different cooperative network metrics than the non-involved category, fitting models with bullying categories as predictors.

Figure 2-7 shows that taking Non-Involved as base category, we find that point estimates of all three bullying categories, compared to Non-Involved students, post a negative effect over Instrength, finding significance at $p < 0.05$ for victims and bully-victims using null hypothesis significance testing, shown in panel A. In B, the same model for Outstrength shows no significant differences between categories and non-involved students. In C, the point estimate for Bully category effect

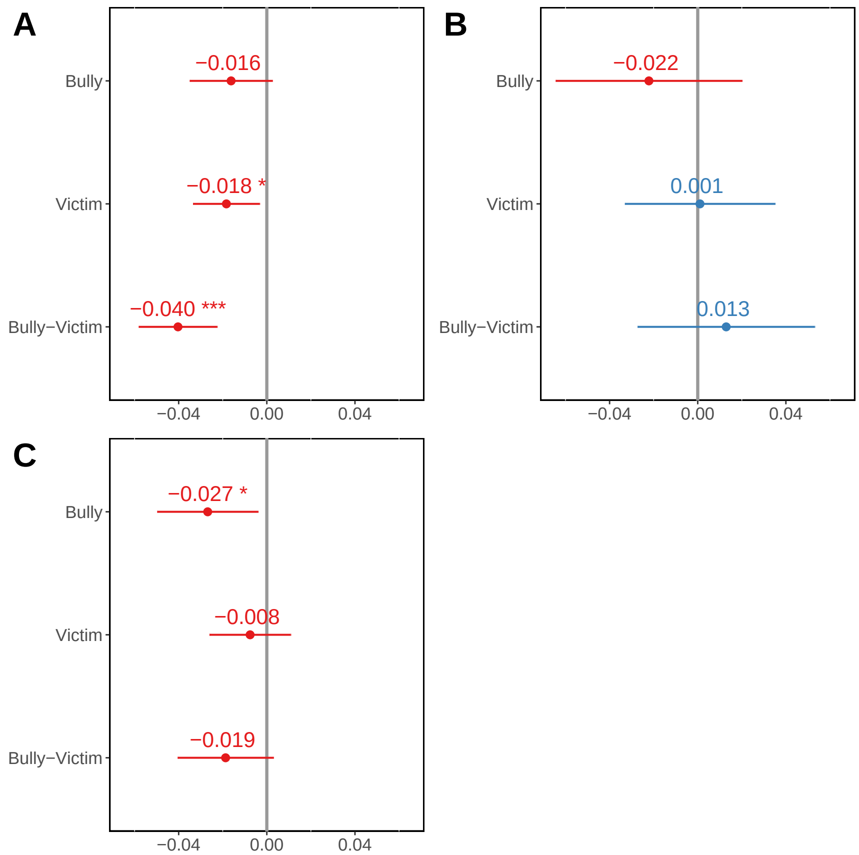

Figure 2-7: **Comparison across bullying categories for topological measures:** Panel A: Instrength, Panel B: Outstrength, Panel C: Reciprocity Strength. Base category: non involved children.

over reciprocity strength was significant and negative.

Three sets of models were built to explore further relations between bullying categories and Instrength, Reciprocity Strength, and Outstrength (the last one was not shown because it did not have any significant relation). All models are standardized and centered following these specific recommendations by Gelman [47]. The model in the first column is the same model from which coefficients were shown in figure 2-7. The second one adds Classroom means for Bully Score and Victim Score. And the third add control variables at the individual and classroom levels.

So, in table 2.1, the Instrength set of models shows that victim and bully/victim negative effects keep sign and significance. The Classroom Victim Score (CVS) variable appears significant with a negative sign and holds after adding controls. We also found a significant negative sign in the comparison of Instrength difference between Grade base category (students in third grade) and 5th

	Instrength	Instrength	Instrength
(Intercept)	0.4987^{***}	0.4964^{***}	0.5365^{***}
	(0.0137)	(0.0127)	(0.0218)
Bully	-0.0162	-0.0162	0.0046
	(0.0096)	(0.0096)	(0.0088)
Victim	-0.0183^{*}	-0.0173^{*}	-0.0187^{**}
	(0.0077)	(0.0077)	(0.0070)
Bully-Victim	-0.0403^{***}	-0.0397^{***}	-0.0267^{**}
	(0.0091)	(0.0091)	(0.0083)
Classroom Mean Bully Score		0.0418	0.0501
		(0.0315)	(0.0366)
Classroom Mean Victim Score		-0.0928^{**}	-0.0814^{*}
		(0.0327)	(0.0342)
Gender (0 = F)			-0.0060
			(0.0065)
Classroom Gender Ratio			-0.0159
			(0.0320)
Attendance			0.0091
			(0.0060)
Classroom Mean Attendance			-0.0307
			(0.0242)
Grade (4th)			-0.0498
			(0.0295)
Grade (5th)			-0.0879^{*}
			(0.0347)
GPA			0.0877^{***}
			(0.0065)
Classroom Mean GPA			-0.0276
			(0.0318)
Num. obs.	1112	1112	1112
Num. groups: clase	47	47	47
Var: clase (Intercept)	0.0080	0.0067	0.0064
Var: Residual	0.0095	0.0095	0.0078

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2.1: In-strength (average weighted in-degree) and bullying categories profiles.

grade, and also a significant positive coefficient with individual GPA and Instrength.

In the Reciprocity Strength set (table 2.2, bully category maintains significance and negative sign when Classroom Bully Score (CBS) and CVS are added to the model. However, this effect vanishes when we control for other variables. The significant terms in this last model are CBS and student's GPA with a positive effect and CVS and 5th Grade category with a negative effect.

	Reciprocity Strength	Reciprocity Strength	Reciprocity Strength
(Intercept)	0.3155***	0.3127***	0.3439***
	(0.0134)	(0.0121)	(0.0213)
Bully	−0.0268*	−0.0269*	−0.0145
	(0.0117)	(0.0117)	(0.0116)
Victim	−0.0076	−0.0058	−0.0073
	(0.0094)	(0.0094)	(0.0092)
Bully-Victim	−0.0187	−0.0178	−0.0113
	(0.0111)	(0.0111)	(0.0109)
Classroom Mean Bully Score		0.0537	0.0727*
		(0.0296)	(0.0356)
Classroom Mean Victim Score		−0.1032***	−0.0962**
		(0.0307)	(0.0332)
Gender (0 = F)			−0.0149
			(0.0086)
Classroom Gender Ratio			−0.0262
			(0.0311)
Attendance			0.0131
			(0.0079)
Classroom Mean Attendance			−0.0230
			(0.0235)
Grade (4th)			−0.0335
			(0.0286)
Grade (5th)			−0.0696*
			(0.0336)
GPA			0.0469***
			(0.0086)
Classroom Mean GPA			−0.0229
			(0.0309)
Num. obs.	1112	1112	1112
Num. groups: clase	47	47	47
Var: clase (Intercept)	0.0071	0.0056	0.0057
Var: Residual	0.0142	0.0142	0.0136

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Table 2.2: Reciprocity-strength (weighted level of mutual tokens exchanged during the participation in the game) and bullying categories profiles

2.4 Discussion

In our experimental setup, every pair of students in a given classroom faced a potential situation where they could exert different cooperation levels with their peers and potentially have mutual benefits if they both honored cooperation, with the risk of being exploited, the temptation of exploiting others, and the certain possibility of end up mutually punishing the other, trapped in a bad nash equilibrium. This setting tried to resemble everyday interactions, where they constantly face many different repeated social dilemmas with other classmates. Thus, we are interpreting the outcome of each dyad as a proxy for the current cooperative disposition toward each other, given the previous interactions between them in similar situations of school life (e.g., the possibility of studying together for a test, sharing food, defend each other, etcetera). Our descriptive results are consistent with this interpretation. The distribution of sent tokens (figure 2-3) is quite variable for students, with the highest frequencies given by extreme values zero and ten, but also a considerable percentage in between them, which in general supports the expected variability of the cooperative dispositions toward different classmates. However, this measure involves both within-students and between-students variations. Going deeper, figure 2-4 helps us isolate within-student variation and keep showing a high variation of sent tokens. Figure 2-5 points to three focal points in dyadic outcomes, which again confirms that even though there is a considerable amount of mutually non-cooperative (or almost-non-cooperative) relations, most of the dyads do not fall into nash equilibrium. We expected this, considering that classmates are not strangers, and many types of relations will influence dyadic behavior away from mutually low cooperative interactions.

It is crucial to make two statements to consider in discussing the remaining results. First, the main result in the literature about bullying and cooperative learning is that a more cooperative environment in the classroom has a consistent effect on lowering bullying and victimization. Second, given the scope and limitations of our data and experimental design, we clearly cannot provide evidence of causality or directionality of the relations we have found. So, to answer our main question about the difference between topological metrics of students self-reported in bullying categories, we will summarize our findings, see how they relate to this main result, and review some general implications and future work.

Self-reported victims and bully/victims received significantly fewer tokens than non-involved classmates. However, we can state that this relationship is robust after controlling other individual and classroom features. These findings might narrow down possible hypotheses to test why and how this consistent effect works. For example, a focused intervention that creates incentives to engage in cooperative behaviors with victims and bully/victims in the classroom, to test whether an increase in received cooperation of victims and bully/victims causes lower future bullying and victimization scores.

Another remarkable outcome is that the Classroom Victimization Score had a negative and significant relation with Instrength that holds with all controls added to the model. This result can be interpreted as a contextual effect, i.e., victims and bully/victims of classes with higher mean victimization score receive even less cooperation than victims and bully/victims from classes with low mean victimization score. This relation was expected and consistent with this major cooperative learning and bullying finding. In that result, there is a directionality that is: more cooperation would cause less victimization. With our method that measures cooperation, one possible way to expand research could be trying the other direction, i.e., testing if we make an intervention that brings mean classroom victimization levels down, would people cooperate more with each other?

In the second set of models, none of the involved categories shows differences in Oustrength with non-involved students. We stated before that there is a considerable amount of variance within-student, but recalling that Oustrength is equivalent to average sent tokens for a student, the absence of a difference between bullying categories shows that this particular point estimate is not very informative. A straightforward suggestion would be to explore more fine-grained metrics that can better characterize this variance within-student.

The third set of models that explains reciprocity displays a negative coefficient for bully category on the simpler models but vanishes when we add all controls. CBS and CVS appear with significant coefficients in the last model, with a positive and negative sign, respectively. CBS positively related to reciprocity might sound paradoxical. A possible explanation would be that the total distribution of students' bully scores is highly right-skewed. However, not all classrooms show this distribution, so classes with higher CBS, holding all else constant, usually have less-skewed, more symmetrical distributions. This would also mean that distance and power imbalance between classmates have decreased, which is a less hierarchical environment that could be a mechanism that explains how higher CBS is associated with higher reciprocity.

What could be some new applications of these findings? • General Interventions: e.g. Cooperative learning: measuring cooperative networks before and after on treated and non-treated groups, to explore how treatment changes cooperative networks and topological measures on individuals. • Targeted interventions: e.g., Giving incentives to collaborate with victims and bully/victims. • Using experimental game theory to measure context-specific behaviors and dispositions in a relational framework. e.g., design a game on a specific social dilemma over behavior in bullying situations. • Combination with other network elicitation methods like sociograms or proximity badges (IoT) to have more precise measurements. • Using these methods to study cooperative networks of small groups in other contexts. e.g., How teams cooperate in different industries like software development, product design, or competitive sports.

Understanding precise mechanisms of the relation between cooperative and bullying behaviors is crucial to improving future generations' learning, engagement, and well-being at our schools. There is plenty of room for major improvements in future work, starting with designing better, actual experiments that make causal inferences. Our game shows impressive preliminary results, but more specific data needs to be collected to state that our measurement is clear and straightforward and that we have thought and ruled out any alternative explanation, effectively narrowing down the options. If we improve this tool and collect more data over more than one time block, we can not just explore but make inferences and try to look for causal evidence over the specific effect of cooperative relations not only on bullying and victimization but on many other important student and classroom outcomes. Given that small-scale cooperative networks are the building blocks for cooperation at large scales, understanding and improving cooperative relations in the classroom is not just important but absolutely necessary as maybe our best chance to face new challenges and global threats such as new pandemics and climate change.

Bibliography

[1] S. Bowles and H. Gintis, *A cooperative species*. Princeton University Press, 2011.

[2] P. Seabright, *The company of strangers*. Princeton University Press, 2010.

[3] R. Sussman, P. Garber, and J. Cheverud, "Importance of cooperation and affiliation in the evolution of primate sociality," *American journal of physical anthropology*, vol. 128, no. 1, p. 84–97, 2005.

[4] S. West, G. Cooper, M. Ghoul, and A. Griffin, "Ten recent insights for our understanding of cooperation," *Nature ecology & evolution*, vol. 5, no. 4, p. 419–430, 2021.

[5] T. Dietz, E. Ostrom, and P. Stern, "The struggle to govern the commons," *Science*, vol. 302, no. 5652, p. 1907–1912, 2003.

[6] A. Woolley, C. Chabris, A. Pentland, N. Hashmi, and T. Malone, "Evidence for a collective intelligence factor in the performance of human groups," *Science*, vol. 330, no. 6004, p. 686–688, 2010.

[7] D. Johnson, R. Johnson, and M. Stanne, "Cooperative learning methods: A meta-analysis," 2000.

[8] A. Thurston, M. Cockerill, and N. Craig, "Using cooperative learning to close the reading attainment gap for students with low literacy levels for grade 8/year 9 students," *International journal of educational research*, vol. 94, p. 1–10, 2019.

[9] M. Van Ryzin and C. Roseth, "Effects of cooperative learning on peer relations, empathy, and bullying in middle school," *Aggressive behavior*, vol. 45, no. 6, p. 643–651, 2019.

[10] D. Olweus, *Aggression in the schools: Bullies and whipping boys*. Hemisphere, 1978.

[11] D. Olweus, "Bullying: What we know and what we can do," *Mental disorder and crime*, p. 353–365, 1993.

[12] D. Olweus, "Aggression and peer acceptance in adolescent boys: Two short-term longitudinal studies of ratings," *Child development*, p. 1301–1313, 1977.

[13] D. Estell, T. Farmer, and B. Cairns, "Bullies and victims in rural african american youth: Behavioral characteristics and social network placement," *Aggressive behavior*, vol. 33, p. 145–159, 2007.

[14] T. W. Farmer, R. A. Petrin, D. L. Robertson, M. W. Fraser, C. M. Hall, S. H. Day, and K. Dadisman, "Peer relations of bullies, bully-victims, and victims: The two social worlds of bullying in second-grade classrooms," *The Elementary School Journal*, vol. 110, no. 3, pp. 364–392, 2010.

[15] D. Schwartz, "Subtypes of victims and aggressors in children's peer groups," *Journal of abnormal child psychology*, vol. 28, p. 181–192, 2000.

[16] T. Vaillancourt, S. Hymel, and P. McDougall, "Bullying is power: Implications for school-based intervention strategies," *Journal of applied school psychology*, vol. 19, p. 157–176, 2003.

[17] P. Rodkin, D. Espelage, and L. Hanish, "A relational framework for understanding bullying: Developmental antecedents and outcomes," *American psychologist*, vol. 70, no. 4, p. 311, 2015.

[18] C. Salmivalli, A. Huttunen, and K. M. Lagerspetz, "Peer networks and bullying in schools," *Scandinavian journal of psychology*, vol. 38, no. 4, pp. 305–312, 1997.

[19] C. Salmivalli, "Bullying and the peer group: A review," *Aggression and violent behavior*, vol. 15, no. 2, p. 112–120, 2010.

[20] S. Perren and F. Alsaker, "Social behavior and peer relationships of victims, bully-victims, and bullies in kindergarten," *Journal of child psychology and psychiatry*, vol. 47, no. 1, p. 45–57, 2006.

[21] A. Kohm, "Childhood bullying and social dilemmas," *Aggressive behavior*, vol. 41, no. 2, p. 97–108, 2015.

[22] B. Schuster, "Rejection and victimization by peers," in *Peer harassment in school: The plight of the vulnerable and victimized* (J. Juvonen and S. Graham, eds.), p. 290–309, The Guilford Press, 2001.

[23] C. Stadtfeld, A. Vörös, T. Elmer, Z. Boda, and I. Raabe, "Integration in emerging social networks explains academic failure and success," *Proceedings of the National Academy of Sciences*, vol. 116, no. 3, p. 792–797, 2019.

[24] C. Candia, M. A. O. Wolf, V. Landaeta, T. Yaikin, C. Monge, C. Hidalgo, and C. Rodriguez-Sickert, "Reciprocity heightens academic performance in elementary school students," 2022.

[25] C. Camerer, "The promise and success of lab-field generalizability in experimental economics: A critical reply to levitt and list," 2011. Available at SSRN 1977749.

[26] D. Karlan, "Using experimental economics to measure social capital and predict financial decisions," *American economic review*, vol. 95, no. 5, p. 1688–1699, 2005.

[27] Y. Algan, Y. Benkler, M. Fuster Morell, and J. Hergueux, "Cooperation in a peer production economy experimental evidence from wikipedia," 2013. Available at SSRN 2843518.

[28] E. Fehr and A. Leibbrandt, "A field study on cooperativeness and impatience in the tragedy of the commons," *Journal of public economics*, vol. 95, no. 9-10, pp. 1144–1155, 2011.

[29] S. Gelcich, R. Guzman, C. Rodríguez-Sickert, J. C. Castilla, and J. C. Cárdenas, "Exploring external validity of common pool resource experiments: insights from artisanal benthic fisheries in chile," *Ecology and Society*, vol. 18, no. 3, 2013.

[30] A. Hopfensitz and J. Miquel-Florensa, "Mill ownership and farmer's cooperative behavior: the case of costa rica coffee farmers," *Journal of institutional economics*, vol. 13, no. 3, p. 623–648, 2017.

[31] M. Rivera-Hechem, R. Guzmán, C. Rodríguez-Sickert, and S. Gelcich, "Effects of experience with access regimes on stewardship behaviors of small-scale fishers," *Conservation biology*, vol. 35, no. 6, p. 1913–1922, 2021.

[32] J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and D. Tracer, "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies," *Behavioral and brain sciences*, vol. 28, no. 6, p. 795–815, 2005.

[33] J. Cárdenas and E. Ostrom, "What do people bring into the game? experiments in the field about cooperation in the commons," *Agricultural systems*, vol. 82, no. 3, p. 307–326, 2004.

[34] C. Rodriguez-Sickert, R. Guzmán, and J. Cárdenas, "Institutions influence preferences: Evidence from a common pool resource experiment," *Journal of economic behavior & organization*, vol. 67, no. 1, p. 215–227, 2008.

[35] A. Pisor, M. Gervais, B. Purzycki, and C. Ross, "Preferences and constraints: the value of economic games for studying human behaviour," *Royal Society open science*, vol. 7, no. 6, p. 192090, 2020.

[36] S. Grueneisen and M. Tomasello, "Children coordinate in a recurrent social dilemma by taking turns and along dominance asymmetries," *Developmental psychology*, vol. 53, no. 2, p. 265, 2017.

[37] T. Yamagishi and K. Cook, “Generalized exchange and social dilemmas,” *Social psychology quarterly*, p. 235–248, 1993.

[38] P. Kollock, “Social dilemmas: The anatomy of cooperation,” *Annual review of sociology*, vol. 24, no. 1, p. 183–214, 1998.

[39] N. Borgers, E. De Leeuw, and J. Hox, “Children as respondents in survey research: Cognitive development and response quality 1,” *Bulletin of Sociological Methodology/Bulletin de méthodologie sociologique*, vol. 66, no. 1, p. 60–75, 2000.

[40] T. Squartini, F. Picciolo, F. Ruzzenenti, and D. Garlaschelli, “Reciprocity of weighted networks,” *Scientific reports*, vol. 3, p. 2729, 2013.

[41] D. Espelage and M. Holt, “Bullying and victimization during early adolescence: Peer influences and psychosocial correlates,” *Journal of emotional abuse*, vol. 2, no. 2-3, p. 123–142, 2001.

[42] D. Espelage, K. Bosworth, and T. Simon, “Examining the social context of bullying behaviors in early adolescence,” *Journal of counseling & development*, vol. 78, no. 3, p. 326–333, 2000.

[43] D. Espelage, M. Holt, and R. Henkel, “Examination of peer-group contextual effects on aggression during early adolescence,” *Child development*, vol. 74, no. 1, p. 205–220, 2003.

[44] J. J. Varela, J. Fábrega, G. Carrillo, M. Benavente, J. Alfaro, and C. Rodríguez-Sickert, “Bullying and subjective well-being: A hierarchical socioeconomical status analysis of chilean adolescents,” *Children and Youth Services Review*, vol. 118, p. 105398, 2020.

[45] Z. Wang, M. Jusup, R. Wang, L. Shi, Y. Iwasa, Y. Moreno, and J. Kurths, “Onymity promotes cooperation in social dilemma experiments,” *Science advances*, vol. 3, no. 3, p. 1601444, 2017.

[46] C. Candia, V. Landaeta-Torres, C. A. Hidalgo, and C. Rodriguez-Sickert, “Strategic reciprocity improves academic performance in public elementary school children,” *arXiv preprint arXiv:1909.11713*, 2019.

[47] A. Gelman, "Scaling regression inputs by dividing by two standard deviations," *Statistics in medicine*, vol. 27, no. 15, p. 2865–2873, 2008.

[48] D. Bates, D. Sarkar, M. Bates, and L. Matrix, "The lme4 package," *R package version*, vol. 2, no. 1, p. 74, 2007.

[49] H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, and H. Yutani, "Welcome to the tidyverse," *Journal of open source software*, vol. 4, no. 43, p. 1686, 2019.

[50] Y. Rosseel, "lavaan: An r package for structural equation modeling," *Journal of statistical software*, vol. 48, p. 1–36, 2012.

[51] D. Lüdecke, "sjplot: Data visualization for statistics in social science," *R package version*, vol. 2, no. 1, 2018.

Appendix A

Chapter 1: Supplementary Material

Supplementary Notes

SM 1.1: Parallel trends in difference-in-differences-like identification strategy

We use a difference in difference-like identification strategy to analyze the impact of reciprocity in academic performance that relies on the condition of the parallel trend for a causal interpretation of results.

It is crucial to consider some socio-economic configurations of Chile to evaluate the plausibility of unobserved time-invariant confounding variables in our difference in differences model specification. First, from elementary to high school, students remain together with their classmates in all classes the whole year, and even the entire school cycle, i. e., students spend together around 8 hours per week, five days a week from March to December each year of the school cycle.

Second, given the academic year starts in March, it completely overlaps with the fiscal year of Chile (which begins on January 1st). Thus, both the socio-economic configurations in Chile and our data collection design, which spans 2017, allow us to assume that classroom and school-level changes are infrequent plausibly. At the individual level, the probability of a student's guardian job changes or student's housing moving is very low within the fiscal year, mainly because these changes usually occur in holidays, between December and March, quite motivated by economic reasons.

Third, we can reasonably assume that some of the variables impacting student engagement also remain time-invariant, such as long programmed absences or failing the school year (this decision is taken by the school at the end of the school year, in December). However, illnesses and family issues are not constrained to systemic configurations, and they directly impact class attendance; that is why we control for class attendance in both observed periods.

Finally, our data comes from vulnerable public schools in a peripheral county of the capital region of Chile. Then we can assume a certain level of homogeneity in household incomes. Moreover, we controlled for the guardian's level of education as a proxy for household income. Therefore, we are also capturing income variations.

SM 1.2: Homophily Index for GPA

Our proposed approach for mapping the social capital network in elementary classrooms unveils the non-existence of homophily between students regarding GPA. We calculate the homophily index [1] as follow:

$$H_r = corr\left(G_i^r, \frac{1}{N^r}\sum_{j}^{N^r} G_j^r\right), \tag{A.1}$$

where G_{ir} represents the GPA of the ego student "i" and G_{jr} represents the GPA of the alter student "j". N indicates the number of students with whom the ego shares the same reciprocity level, r. The superindex r represents the level of reciprocity in the interaction. Those levels are: i) Low reciprocity (0.0) includes the interactions where egos and alters send fewer tokens than their average amount sent to the rest of their class. ii) Positive asymmetry (0.1) includes the interactions where alters send more tokens than their average sent tokens and egos send fewer tokens than their average sent amount, regarding their class. iii) Negative asymmetry (1.0) includes the interactions where egos send more tokens than their average sent tokens and alters send fewer tokens than their average sent amount, regarding their class. iv) High reciprocity (1.1) includes the interactions where egos and alters send more tokens than their average amount sent to the rest of their class.

Fig. A-1 shows that in the "Cooperation Network" most of the classes have Homophily Indexes non-significantly different from zero, with both GPA before and after measuring. This finding is interesting because it has been shown that GPA modulates social relationship in older students–High School, undergrads, and grad students–[1–3], however, here we provide evidence on this behavioral pattern does not manifest in elementary school students. Thus, interventions to promote and boost cooperative social relationships could be possible to improve academic performance, by exposing students with low GPA to the idea flow and information of the students with high GPA [4].

These results implies that cooperative interactions in children do not seem to be driven by academic homophily.

Table A.1: Homophily Index of Cooperation Networks on GPA before and after measuring, by Class

Class ID	Homophily Index GPA (Before)	Homophily Index GPA (After)
10	-0.1947	0.0042
11	-0.1993	-0.3993
12	0.2509	0.5652 **
13	-0.2539	-0.3811
14	0.8272 ***	0.7763 ***
15	0.5044	0.0851
16	-0.2429	-0.2542
17	-0.2446	-0.0952
18	0.0499	-0.2112
19	0.0986	-0.1119
20	-0.0348	0.0633
21	0.0708	0.1366
23	0.0893	0.5436 ***
24	-0.1848	0.2128
25	-0.311	-0.456 **
26	-0.0834	-0.3109
27	-0.1472	-0.1304
28	0.1729	0.462 **
29	-0.5273 **	-0.2093
30	-0.3425	-0.4028
31	-0.2936	0.1011
32	-0.2142	-0.1654
33	0.3756 *	-0.1863
34	-0.8316 ***	-0.2617
35	0.5746 *	-0.2938
38	-0.1215	0.2298
39	-0.2444	-0.48 *
40	0.3102	0.6662 ***
41	0.0607	0.1991
42	-0.5799 *	-0.3401
43	0.045	-0.4457
44	-0.1497	-0.0175
45	-0.09	-0.5937
46	-0.5 **	-0.305
47	-0.6942 *	-0.6366
48	-0.8405 **	-0.6899
49	-0.3657	-0.1622
50	0.2362	0.2207
51	-0.0339	-0.1241
52	0.2206	0.4032 *
55	-0.0127	0.0591
56	-0.0086	0.4444 *
57	-0.454 **	-0.3179
58	-0.1344	0.0492
59	0.1459	-0.5666 **

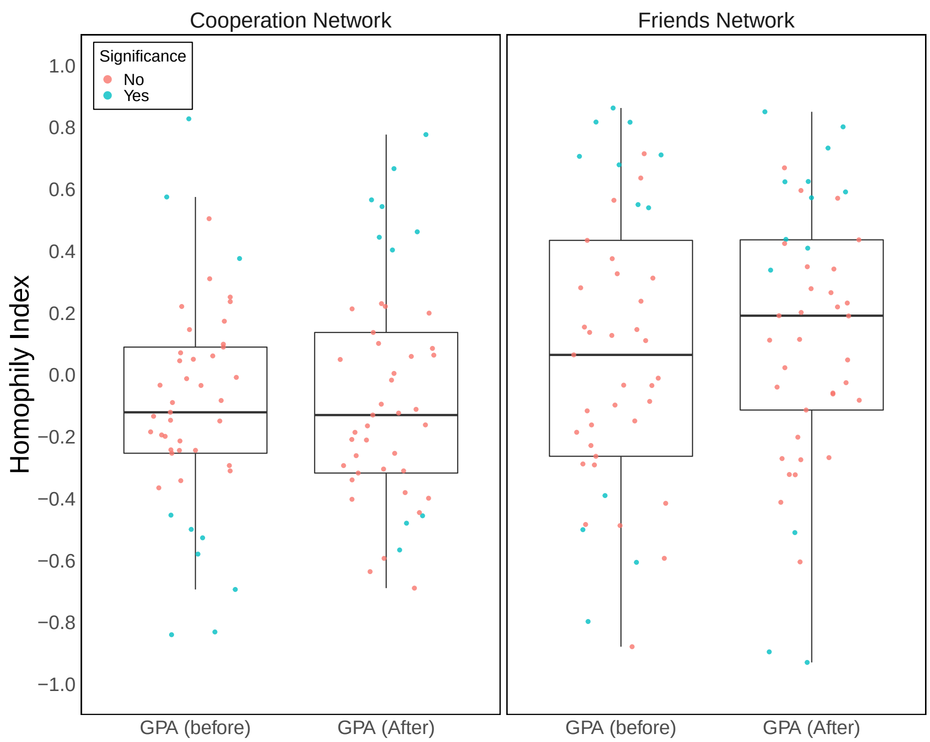

Figure A-1: Homophily index according to GPA in elementary school students. The homophily index for high reciprocity is non-significantly different from zero, which means that elementary students do not build their cooperative disposition based on the GPA of their peers. The significance is given by color, and it represents a p-value under 0.1. Most of the rest of the homophily coefficients are low if we compare them with the literature for older students [1], where the initial level of homophily is greater than 0.2 and evolves until 0.4.

Supplementary Tables

SM 2.1: Classrooms descriptions

Table A.2: Class-level description.

Class ID	School ID	Age (Mean)	Age (SD)	Female (%)	Guardian education Secondary school completed (%)	GPA (Mean)	Class size	Attendance (Mean %)	Cooperation (Mean tokens)
1	1	10.03	0.57	57.6	48.5	6	33	89.3	3.97
2	2	9.09	0.51	41.7	45.8	6	24	91.9	5.59
3	2	10.17	0.51	53.3	56.7	5.9	30	90.2	4.76
4	2	10.87	0.66	43.5	56.5	5.9	23	89.2	4.39
5	3	9.15	0.99	30.8	30.8	5.7	13	89.2	6.09
6	3	9.05	0.77	56.2	50	5.6	16	93.3	5
7	3	10.39	0.9	39.1	34.8	5.8	23	91.6	4.86
8	3	10.91	0.7	33.3	50	5.7	18	93.7	4.3
9	3	11.46	1.3	52.6	52.6	5.2	19	88.9	7.14
10	4	9.17	0.64	37.5	50	6	8	94.6	7.23
11	4	10.01	0.95	42.9	28.6	6	7	97.5	4.88
12	4	11.06	0.77	42.9	38.1	6.1	21	88.9	5.1
13	5	10.05	0.72	55	30	6.6	20	90.3	5.04
14	5	10.18	0.78	58.8	35.3	6	17	91	4.21
15	5	11.22	0.63	50	30.8	5.8	26	93.7	3.54
16	6	9.41	1.05	66.7	55.6	5.9	9	92	6.25
17	6	10.3	0.95	33.3	50	5.5	12	88.3	5.92
18	6	11.66	2.56	35.7	64.3	5.8	14	89.6	4.06
19	7	9.17	0.74	73.3	26.7	6.1	15	95.2	4.38
20	7	10.3	0.84	47.6	28.6	6	21	95.3	5.9
21	7	11.24	0.73	33.3	44.4	5.8	18	94.4	4.63
22	8	9.2	0.64	46.2	43.6	5.7	39	89.8	5.08
23	8	10.11	0.64	63.6	42.4	5.9	33	92.4	4.06
24	8	11.12	0.55	48.4	45.2	5.4	31	91.2	4.34
25	9	8.98	0.63	100	44	6.2	25	88.8	6.78
26	9	9.87	0.76	100	35.7	6	14	83.7	5.32
27	9	10.73	0.65	100	50	6	14	89.4	5.31
28	10	9.07	0.92	36.4	22.7	5.7	22	86.8	3.79
29	10	9.38	1.12	61.5	30.8	5.7	13	91.3	7.22
30	10	10.12	0.59	45.5	22.7	5.8	22	85.2	3.48
31	10	10.35	0.95	50	22.7	5.6	22	84.6	4.05
32	10	11.44	1.08	53.6	25	5.7	28	88.3	3.79
33	11	8.8	0.52	0	50	5.7	14	90.7	6.81
34	11	10.3	0.43	0	75	6	12	94.3	3.92
35	11	11	0.67	0	50	6.2	10	94.2	3.09
36	12	9.47	0.83	58.3	25	5.8	12	90.8	5.59
37	12	10.2	0.88	20	30	5.6	10	87.9	4.39
38	12	11.65	1.11	35.3	47.1	5.5	17	89.3	3.2
39	13	9.92	0.5	100	69.2	6.1	26	94.2	3.62
40	13	9.97	0.49	100	38.9	5.8	18	91.8	4.67
41	13	9.8	0.61	100	70	6.2	20	94.8	5.29
42	14	8.91	0.61	100	72.2	6.3	18	96	5.33
43	14	8.97	0.69	100	50	6	18	91.6	5.6
44	14	10.04	0.56	100	63.2	6	19	91	5.11
45	14	11.7	1.64	100	73.3	6.2	15	96.2	4.09

Supplementary Figures

SM 3.1: Correlation Matrix

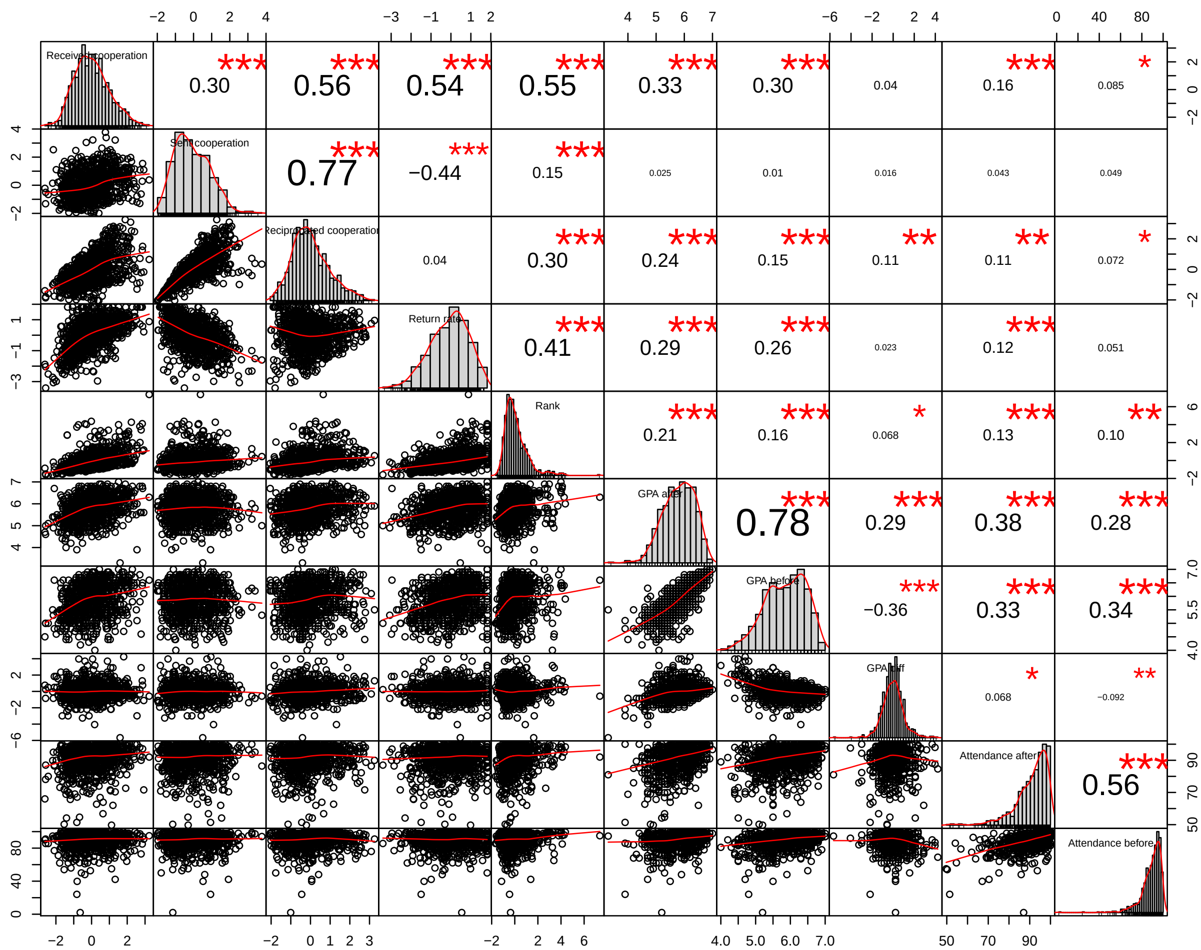

Figure A-2: Correlations between all collected, measured, and built variables.

SM 3.2: Sent tokens by students

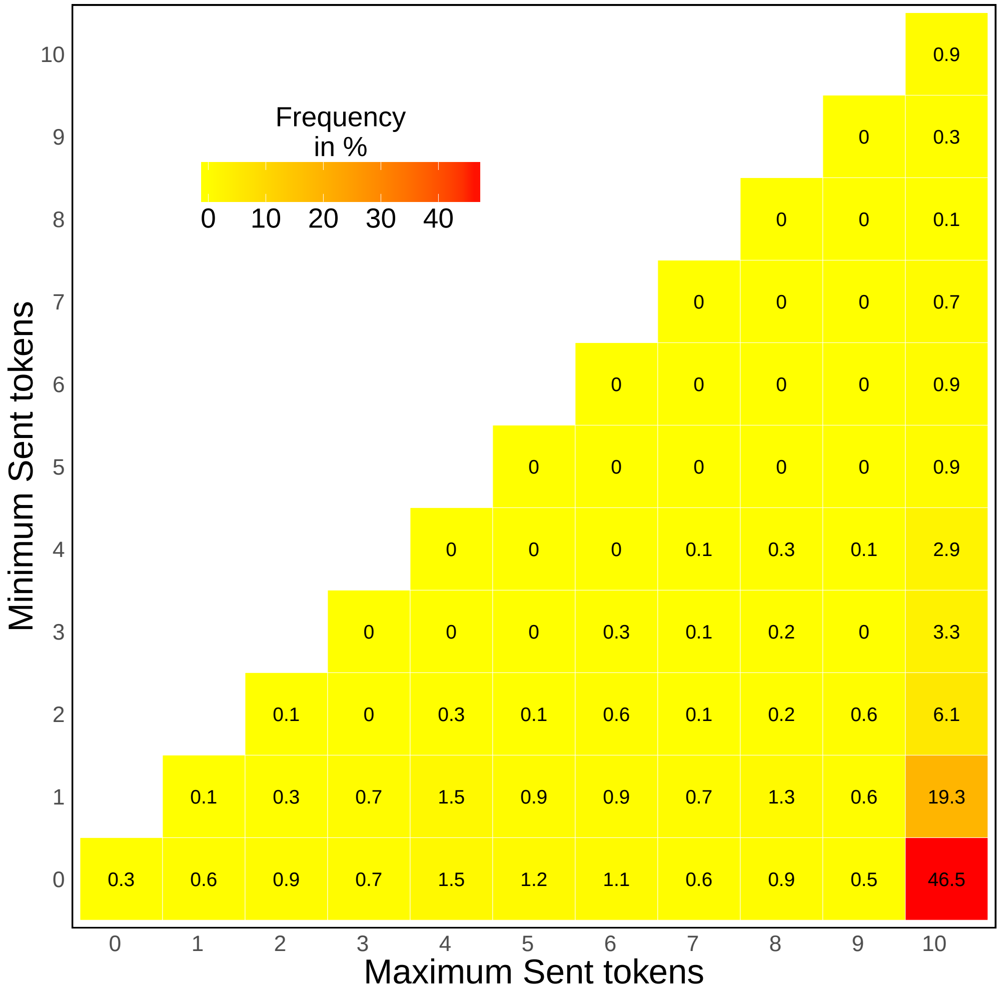

Figure A-3: Heatmap for maximum and minimum sending. At the bottom-right corner, we observe a cluster ($\approx 65\%$ of students, red/orange blocks) indicating that students send tokens in the whole range of possibilities. History between colleagues matters.

SM 3.3: Explanatory Power of the Model

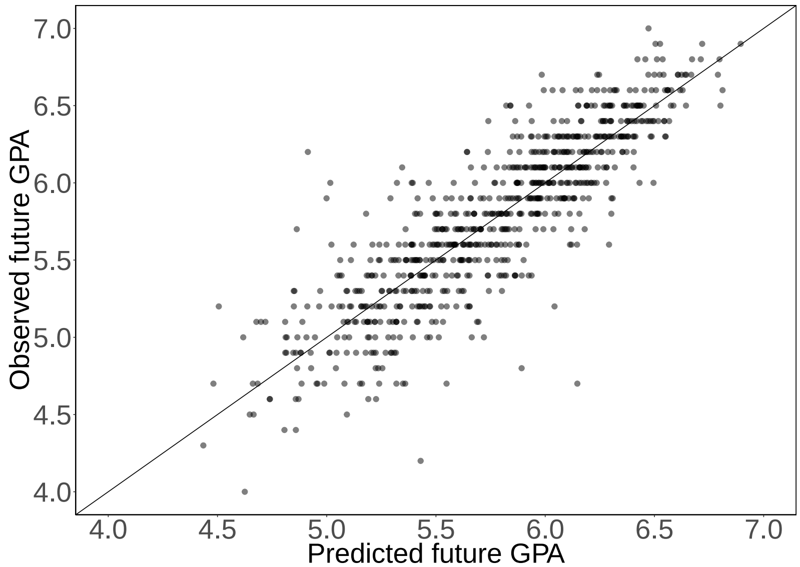

Figure A-4: Observed v/s predicted GPA (table 2 model 3)

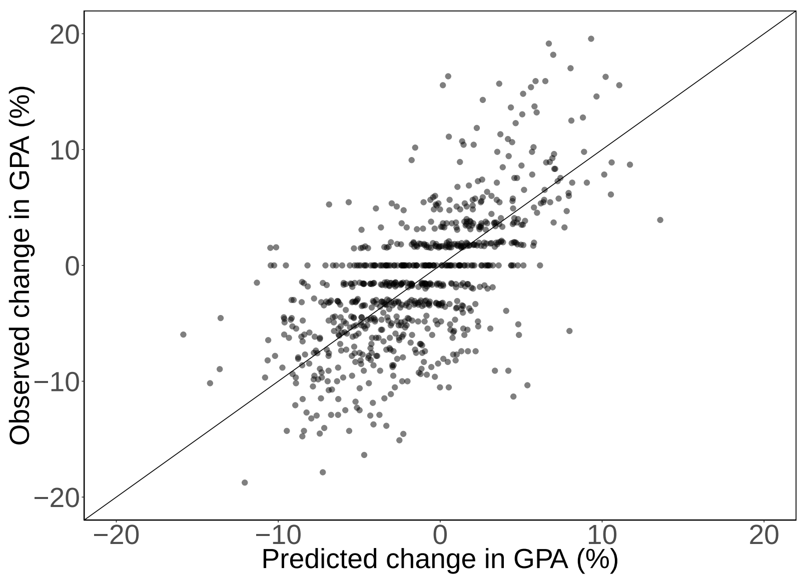

Figure A-5: Observed v/s predicted change in GPA .

SM 3.4: Reciprocity and confounders

	Dependent variable:						
	Reciprocity (z-score)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sent cooperation (z-score)	0.712***		0.706***			0.705***	0.705***
	(0.02)		(0.01)			(0.02)	(0.02)
Rank (z-score)		0.449***	0.425***			0.343***	0.343***
		(0.05)	(0.02)			(0.02)	(0.02)
Grades (before measuring)				0.255***	0.207***	0.158***	0.158***
				(0.05)	(0.05)	(0.03)	(0.03)
Attendance (%)					0.008**	0.003*	0.003*
					(0.00)	(0.00)	(0.00)
Tutor comp. sec. school (yes)					0.018	0.037	0.037
					(0.06)	(0.03)	(0.03)
Sex (Male)					-0.166***	-0.119***	-0.119***
					(0.06)	(0.03)	(0.03)
Constant	-0.007	-0.003	-0.007	-1.499***	-1.883***	-1.172***	-1.172***
	(0.02)	(0.02)	(0.01)	(0.29)	(0.38)	(0.19)	(0.19)
Fixed effects	Class-group	Class-group	Class-group	Class-group	Class-group	Class-group	Class-group
Observations	859	859	859	859	771	771	771
R-squared	0.815	0.507	0.868	0.465	0.489	0.883	0.883
Adjusted R-squared	0.804	0.480	0.861	0.435	0.455	0.875	0.875
R-squared within	0.665	0.108	0.761	0.032	0.053	0.784	0.784
F Statistics	1610.971	98.361	1294.442	26.729	10.085	435.206	435.206
Note:							*p<0.1; **p<0.05; ***p<0.01

Table A.3: OLS Estimation of Reciprocity (class-group fixed effects)

SM 3.5: Robustness of difference in difference estimation

	Dependent variable:										
	Reciprocity (Diff-in-Diff)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Reciprocity * Time	0.039***	0.038***									
	(0.01)	(0.01)									
Time	-0.081***	-0.094***	-0.086***	-0.087***	-0.080***	-0.075***	-0.069***	-0.125***	-0.113***	-0.114***	-0.105***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Attendance (%)		0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bottom 10% Reciprocity * Time			-0.089**								
			(0.04)								
Bottom 20% Reciprocity * Time				-0.037							
				(0.03)							
Bottom 30% Reciprocity * Time					-0.050*						
					(0.03)						
Bottom 40% Reciprocity * Time						-0.048*					
						(0.03)					
Bottom 50% Reciprocity * Time							-0.051**				
							(0.03)				
Top 40% Reciprocity * Time								0.075***			
								(0.03)			
Top 30% Reciprocity * Time									0.060**		
									(0.03)		
Top 20% Reciprocity * Time										0.100***	
										(0.03)	
Top 10% Reciprocity * Time											0.094**
											(0.04)
Fixed effects	Individual	Individual	Individual	Individual	Individual	Individual	Individual	Individual	Individual	Individual	Individual
Observations	1710	1710	1710	1710	1710	1710	1710	1710	1710	1710	1710
R-squared	0.894	0.897	0.896	0.896	0.896	0.896	0.896	0.897	0.896	0.897	0.896
Adjusted R-squared	0.787	0.793	0.791	0.791	0.791	0.791	0.791	0.792	0.792	0.793	0.792
R-squared within	0.054	0.082	0.077	0.073	0.075	0.076	0.076	0.081	0.077	0.083	0.078
F Statistics	24.535	25.217	23.584	22.506	23.164	23.254	23.434	24.993	23.727	25.604	23.900

Note: *p<0.1; **p<0.05; ***p<0.01

Table A.4: Diff-Diff Estimation of Reciprocity (individual fixed effects)

Bibliography

[1] I. Smirnov and S. Thurner, "Formation of homophily in academic performance: Students change their friends rather than performance," *PLOS ONE*, vol. 12, p. e0183473, aug 2017.

[2] V. Kassarnig, E. Mones, A. Bjerre-Nielsen, P. Sapiezynski, D. Dreyer Lassen, and S. Lehmann, "Academic performance and behavioral patterns," *EPJ Data Science*, vol. 7, p. 10, dec 2018.

[3] C. Candia, S. Encarnação, and F. L. Pinheiro, "The higher education space: connecting degree programs from individuals' choices," *EPJ Data Science*, vol. 8, no. 1, p. 39, 2019.

[4] A. Pentland, *Social Physics: How Social Networks Can Make Us Smarter*. Penguin Books, 2015.

Appendix B

Chapter 2: Supplementary Material

Descriptive Statistics

Table 1: Descriptive Statistics

	Variables	M	SD	Range	N
4	Instrength	0.479	0.133	0.8 (0.08-0.88)	1137
5	Outstrength	0.479	0.236	1 (0-1)	1137
6	Reciprocated Strength	0.300	0.146	0.81 (0-0.81)	1137
2	Bully Score	1.699	0.598	3 (1-4)	1137
7	Victim Score	2.375	0.905	3 (1-4)	1137
3	GPA	5.760	0.555	3.7 (3.3-7)	1112
1	School Attendance	90.838	8.556	94.23 (5.77-100)	1114

Figure B-1: Mean and Standard Deviation of main numeric variables

Zero-Order Correlations

	Instrength	*Outstrength*	*Reciprocated Strength*	*Bully Score*	*Victim Score*	*GPA*	*School Attendance*
Instrength	-						
Outstrength	0.272***	-					
Reciprocated Strength	0.659***	0.842***	-				
Bully Score	-0.093**	-0.034	-0.085**	-			
Victim Score	-0.176***	-0.028	-0.104***	0.332***	-		
GPA	0.369***	0.049	0.219***	-0.226***	-0.077**	-	
School Attendance	0.120***	0.033	0.086**	-0.020	0.038	0.334***	-

Computed correlation used pearson-method with listwise-deletion.

Figure B-2: Zero order correlation matrix

Bullying Categories	*Gender* Female	Male	***Total***
Not Identified	345 53.2 %	232 47.4 %	577 50.7 %
Bully	63 9.7 %	79 16.2 %	142 12.5 %
Victim	152 23.5 %	104 21.3 %	256 22.5 %
Bully-Victim	88 13.6 %	74 15.1 %	162 14.2 %
Total	648 100 %	489 100 %	1137 100 %

χ^2=12.145 · *df*=3 · *Cramer's V*=0.103 · *p*=0.007

Figure B-3: Contingency table of Bullying Categories and Gender

Confirmatory Factor Analysis and Item Reliability on SRBS and SRVS

Given that University of Illinois self-report survey [1] was previously validated using EFA and found three constructs, Bully, Victim and Fight, we're performing a three-factor CFA to confirm a three-factor solution. After confirming this factor model, we're aggregating quantitative scores for each construct more confidently. CFA was run using the R package lavaan [3]. Taking into account that likert scores are ordinal and usually nonnormal, the robust maximum-likelihood estimation (MLR) was used to extract the variances from the data. Following [1], we specified a three-factor CFA, with questions 1-9 (10–18 in this coded dataset) representing Bully Scale, items 10-14 (19–23) representing Fight Scale, and finally items 15-18 (24-27) for Victimization Scale. Multiple fit indices (chi-square value from robust MLR [MLR χ^2]; comparative fit index [CFI]; the root-mean-square error of approximation [RMSEA]; and the standardized root-mean-square residual [SRMR]) were consulted to evaluate model fit. Consistent with the recommendations by Hu and Bentler [2], the following criteria were used to evaluate the adequacy of the models: $CFI > 0.95$, $SRMR < 0.08$, and $RMSEA < 0.06$. Coefficient alpha was computed based on the model results and used to assess reliability. Values > 0.70 were considered acceptable. CFA output using lavaan was the following:

Results from the three-factor CFA indicated that, in our population, the data did support the model specified. The chi-square test of model fit was significant ($\chi^2 = 5547.219, df = 153, p < 0.000$) and that would be a signal that the factor model is not supported, but this test is known to be sensitive to minor model misspecification with large sample sizes (n = completar) (reference). On the other hand, additional model fit criteria indicated that the data did support the model specified. Robust CFI is higher than 0.95 cutoff. Robust RMSEA point estimation was 0.041, and 90% CI is (0.035 - 0.046) so we're definitely below the 0.06 cutoff. SRMR was 0.041, which also suggests a good fit. Item analysis and Cronbach's alpha calculations were explored with package sjPlot. We also cross-checked with function alpha from package psych and results are consistent.

```
summary(C2f_fit, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
lavaan 0.6-10 ended normally after 35 iterations

 Estimator                          ML
 Optimization method                NLMINB
 Number of model parameters         39

 Number of observations             1137

 Model Test User Model:
                        Standard    Robust
 Test Statistic          419.815   322.373
 Degrees of freedom          132       132
 P-value (Chi-square)      0.000     0.000
 Scaling correction factor           1.302
   Yuan-Bentler correction (Mplus variant)

Model Test Baseline Model:

 Test statistic         6848.206  5144.316
 Degrees of freedom          153       153
 P-value                   0.000     0.000
 Scaling correction factor           1.331

User Model versus Baseline Model:

 Comparative Fit Index (CFI)     0.957     0.962
 Tucker-Lewis Index (TLI)        0.950     0.956

 Robust Comparative Fit Index (CFI)        0.963
 Robust Tucker-Lewis Index (TLI)           0.957

Loglikelihood and Information Criteria:

 Loglikelihood user model (H0)       -25301.946  -25301.946
 Scaling correction factor                1.261
   for the MLR correction
 Loglikelihood unrestricted model (H1)  -25092.038  -25092.038
 Scaling correction factor                1.293
   for the MLR correction

 Akaike (AIC)                      50681.892   50681.892
 Bayesian (BIC)                    50878.301   50878.301
 Sample-size adjusted Bayesian (BIC)  50754.426   50754.426

Root Mean Square Error of Approximation:

 RMSEA                              0.044     0.036
 90 Percent confidence interval - lower0.039     0.031
 90 Percent confidence interval - upper0.049     0.040
 P-value RMSEA <= 0.05              0.985     1.000

 Robust RMSEA                                 0.041
 90 Percent confidence interval - lower 90    0.035
 Percent confidence interval - upper          0.046

Standardized Root Mean Square Residual:
```

Figure B-4

Bully

Missings	*Mean*	*SD*	*Skew*	*Item Difficulty*	*Item Discrimination*	*α if deleted*
0.00 %	2.2	0.99	0.11	0.55	0.61	0.80
0.00 %	1.65	0.92	1.19	0.33	0.56	0.81
0.00 %	1.75	1	0.99	0.35	0.57	0.81
0.00 %	1.55	0.84	1.31	0.39	0.40	0.82
0.00 %	1.53	0.91	1.58	0.31	0.50	0.81
0.00 %	1.51	0.87	1.57	0.30	0.49	0.81
0.00 %	1.7	0.92	0.97	0.43	0.49	0.81
0.00 %	1.63	0.91	1.23	0.33	0.61	0.80
0.00 %	1.8	0.97	0.87	0.36	0.54	0.81

Mean inter-item-correlation=0.347 · Cronbach's α=0.828

Fight

Missings	*Mean*	*SD*	*Skew*	*Item Difficulty*	*Item Discrimination*	*α if deleted*
0.00 %	1.63	0.94	1.22	0.33	0.64	0.74
0.00 %	1.53	0.91	1.53	0.31	0.58	0.76
0.00 %	1.44	0.81	1.76	0.29	0.53	0.77
0.00 %	2.25	1.16	0.26	0.45	0.57	0.77
0.00 %	1.61	0.95	1.32	0.32	0.59	0.75

Mean inter-item-correlation=0.445 · Cronbach's α=0.796

Victim

Missings	*Mean*	*SD*	*Skew*	*Item Difficulty*	*Item Discrimination*	*α if deleted*
0.00 %	2.62	1.11	-0.2	0.52	0.65	0.73
0.00 %	2.42	1.12	0.05	0.48	0.68	0.72
0.00 %	2.32	1.21	0.19	0.58	0.60	0.76
0.00 %	2.16	1.12	0.35	0.43	0.54	0.79

Mean inter-item-correlation=0.504 · Cronbach's α=0.802

	Component 1	*Component 2*	*Component 3*
Component 1	α=0.828		
Component 2	0.726 *(<.001)*	α=0.796	
Component 3	0.341 *(<.001)*	0.310 *(<.001)*	α=0.802
Computed correlation used pearson-method with listwise-deletion.			

Figure B-6

Bibliography

[1] Espelage, D. and Holt, M. (2001). Bullying and victimization during early adolescence: Peer influences and psychosocial correlates. *Journal of emotional abuse*, 2(2-3):123–142.

[2] Hu, L.-t. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1):1–55.

[3] Rosseel, Y. (2012). lavaan: An r package for structural equation modeling. *Journal of statistical software*, 48:1–36.

List of Figures

1-1 Experimental game. A) Students play a social dilemma; example of a dyadic interaction: i) Both players are endowed with 10 (ten) tokens. ii) Simultaneously, player A decides to send 3 (three) tokens, and player B decides to send 2 (two) tokens. iii) Sent tokens are doubled. iv) Player A receives 4 (four) tokens and player B receives 6 (six) tokens. B) A students' network for a single classroom. For the sake of visualization, edges correspond only to fully cooperative interactions. . . . 18

1-2 Patterns of sent and received tokens across all peer interactions. A) Bivariate distribution of sent tokens. We observe three peaks describing the most commonly used strategies: I) Social optimum (top-right). II) Nash equilibrium (bottom-left). III) Asymmetric exchange (bottom-right). B) Univariate distribution for sent tokens. 21

2-1 The Dyadic Social Dilemma. Example: I) Both players are endowed with 10 (ten) tokens. B) Simultaneously, Player A decides to send 3 tokens, and Player B decides to send 2 tokens. III) Sent tokens are doubled. Player A receives 4 tokens and second player receives 6 tokens. Note: Reprinted from [46] 48

2-2 Example from a hypothetical class with $N = 4$ students. A) Tablet computer platform. B) Adjacency matrix where w_{ij} represents the tokens sent from student i to student j in the dyadic game. C) Cooperative weighted network that graphically represents the information displayed in the adjacency matrix. The width of the edge represents the amount of sent tokens and the node's size, the number of tokens each student ended up with at the end of the game. Node color indicates the bullying category assigned to the student. 49

2-3 **Individual behavior in the dyadic social dilemma:** Histogram of sent tokens (individual level). 50

2-4 **Histogram of min-max sent tokens by students** 51

2-5 **Dyadic behavior in the social dilemma:** Histogram of experimental outcomes at the dyadic level over all sample. 52

2-6 **Classroom reciprocity network:** Red nodes correspond to bullies, purple nodes to bully/victims, light blue to victims, and green nodes correspond to non-involved. Linked nodes correspond to relationships in which both members of the dyad sent more tokens to their partners than what they sent on average to the rest of the class members, i.e., they have a relation of mutual cooperation. In this case, victim and bully/victims are mostly located on the periphery of this network. 53

2-7 **Comparison across bullying categories for topological measures:** Panel A: Instrength, Panel B: Outstrength, Panel C: Reciprocity Strength. Base category: non involved children. 54

A-1 Homophily index according to GPA in elementary school students. The homophily index for high reciprocity is non-significantly different from zero, which means that elementary students do not build their cooperative disposition based on the GPA of their peers. The significance is given by color, and it represents a p-value under 0.1. Most of the rest of the homophily coefficients are low if we compare them with the literature for older students [1], where the initial level of homophily is greater than 0.2 and evolves until 0.4. 71

A-2 Correlations between all collected, measured, and built variables. 73

A-3 Heatmap for maximum and minimum sending. At the bottom-right corner, we observe a cluster ($\approx 65\%$ of students, red/orange blocks) indicating that students send tokens in the whole range of possibilities. History between colleagues matters. 74

A-4 Observed v/s predicted GPA (table 2 model 3) 75

A-5 Observed v/s predicted change in GPA 75

B-1 Mean and Standard Deviation of main numeric variables 79

B-2 Zero order correlation matrix 80

B-3 Contingency table of Bullying Categories and Gender 80

B-4 82

B-5 83

B-6 84

List of Tables

1.1 Network measures. w_{ij} is the number of tokens sent from i to j, d represents a dumping factor ($d = 0.85$ following [78]), and N is the number of students in each classroom 19

1.2 OLS regressions for students' GPA after measuring cooperation. Note that these models show results for 769 students because we do not have the data for Tutor complete secondary school for 86 students. 24

1.3 Difference-in-differences-like estimation of GPA. Models 1 and 2 differ in the control variable "class attendance". In Models 3 and 4, we built a dummy variable for reciprocated cooperation. It takes value 1 if the reciprocated cooperation level for individual i is in the top 20%, and it takes value 0 otherwise. 25

2.1 In-strength (average weighted in-degree) and bullying categories profiles. 55

2.2 Reciprocity-strength (weighted level of mutual tokens exchanged during the participation in the game) and bullying categories profiles 57

A.1 Homophily Index of Cooperation Networks on GPA before and after measuring, by Class 70

A.2 Class-level description. 72

A.3 OLS Estimation of Reciprocity (class-group fixed effects) 76

A.4 Diff-Diff Estimation of Reciprocity (individual fixed effects) 77