Academic achievement helps coordination on mutually advantageous outcomes *

Isabelle Brocas
University of Southern California
IAST and CEPR

Juan D. Carrillo
University of Southern California
IAST and CEPR

Abstract

This study examines the relationship between academic achievement and strategic ability to coordinate among middle school students. We designed an experimental framework using repeated asymmetric Battle of the Sexes and Hawk-Dove games, to explore how cognitive and social skills related to academic success influence behavior. A total of 132 students participated, divided into groups of high and low academic achievers based on their performance at school. Our results show that, on average, high achievers coordinate better on equilibrium outcomes with simple but effective strategies and obtain higher payoffs compared to low achievers. This indicates that academic success may reflect broader cognitive abilities–such as strategic thinking, anticipation of others' choices, and cooperation–crucial for navigating real-world interactions. However, we notice also substantial heterogeneity within groups. Finally, performance in pairs with one high and one low achiever is intermediate but closer to the level of high achievers, suggesting potential peer learning effects and the educational value of mixed groups to promote guidance and joint improvements. These findings emphasize the role of long-term learning in developing cognitive skills that facilitate cooperation and strategic interaction in complex environments.

Keywords: developmental game theory, coordination, repeated games.

Significance Statement

This study explores how academic achievement relates to strategic coordination among middle school students. Using classical coordination games, we examined how cognitive and social skills linked to academic success impact behavior. We divided our 132 participants into high and low achievers. Results showed that high achievers coordinate better, using simple strategies and earning higher payoffs than low achievers. This suggests that academic success reflects broader cognitive abilities, such as strategic thinking and cooperation. Additionally, mixed pairs (with one high and one low achiever) performed closer to high achievers, hinting at peer learning effects and the value of diverse groupings for skill development.

*We are grateful to members of the Los Angeles Behavioral Economics Laboratory (LABEL), especially Ulysses Rios and Mónica Vasco, for their insights and comments and to Thomas Starr King Middle School, especially Adriana Díaz, for their help and support. We gratefully acknowledge financial support of the National Science Foundation (NSF) under the grant SES-2315770, and IAST funding from the French National Research Agency (ANR) under the grant ANR-17-EURE-0010. Address for correspondence: Department of Economics, University of Southern California, 3620 S. Vermont Ave., Los Angeles, CA 90089, USA, <brocas@usc.edu> and <juandc@usc.edu>.

1 Introduction

In game theory, coordination refers to the process by which players align their strategies to achieve a mutually beneficial outcome. This is especially important in games with multiple equilibria—situations where different outcomes are possible where no player has an incentive to change their strategy on their own (Camerer, 2003; Cooper and Weber, 2020).

Coordination games are highly relevant for modeling real-world challenges such as compromise in relationships, decision-making in team projects, dynamic partnerships, and conflicts over shared resources. A substantial body of experimental research has highlighted differences in the complexity of coordination mechanisms, often linked to variations in personality traits (Proto et al., 2019), cultural backgrounds (Jackson and Xing, 2014; Peryman and Kelsey, 2021) and identity of the peers involved (Eckel and Wilson, 2007; Lambrecht et al., 2024). However, developmental aspects of these abilities, particularly during adolescence—a critical period for acquiring cognitive and social skills—remain under-explored.

Recent research from our group indicates that the ability to coordinate develops throughout childhood and adolescence (Brocas and Carrillo, 2023). Adolescence is widely recognized as a pivotal period for developing decision-making styles that often persist into adulthood. Differences in coordination mechanisms seen in adults may stem from variations in intellectual ability, emotional intelligence, or the capacity to read social cues—skills that mature during adolescence and blend into a sophisticated composite ability essential for strategic interaction. Does academic achievement during this formative stage serve as a reflection of this multifaceted skill? Can it reliably predict performance in complex coordination tasks? Despite its relevance, little research has explored how long-term learning ability, captured by academic success, influences adolescents' ability to navigate strategic environments.

In this study, we investigate the link between academic achievement and strategic coordination abilities in adolescents, focusing on middle school students (ages 11 to 14). Academic success, as a composite measure of cognitive and social skills developed through formal education, may reflect broader abilities crucial for strategic interactions. We hypothesize that high academic achievers are better equipped to anticipate others' actions, employ effective strategies, and obtain higher payoffs in repeated coordination games. Also, pairs involving one high and one low achiever may demonstrate intermediate outcomes, suggesting potential peer learning effects. Finally, low achievers may struggle with these tasks, reflecting challenges in problem-solving, attention, and cooperative behaviors. To explore this idea, we design a cooperative experimental paradigm and test the behavior of adolescents with varying levels of academic success.

A main factor in coordination is whether the game is played once or repeatedly. In a one-shot game without communication, coordination can be challenging. In many real-world scenarios, however, individuals often interact multiple times with the same people. Repeated play allows them to learn from past experiences and signal their intentions to try and resolve potential miscoordinations. Overall, while one-shot games are useful for studying how norms influence coordination, repeated interactions involve identification of patterns from previous play and strategic signaling—factors we believe are more closely linked to academic accomplishments. For this reason, we focus on a repeated coordination framework.

A key finding in related studies is the remarkable ability of adults to achieve long-term coordination on fair and Pareto-efficient equilibria in games like the Stag Hunt and the symmetric Battle of the Sexes, where coordination in one-shot versions is rare (McKelvey and Palfrey, 2001; Ioannou and Romero, 2014). This pattern of behavior develops gradually during childhood and solidifies in adolescence (Brocas and Carrillo, 2023). In both cases, adolescents and adults alike often recognize the benefits of a cooperative approach. Intuitive ideas of fairness and efficiency are less effective in the Hawk-Dove (or chicken) game, which involves balancing aggression and avoidance, and in asymmetric versions of the Battle of the Sexes, where long-term fairness requires more complex turn-taking strategies. We decided to employ these two arguably more nuanced and complex games because we anticipated that differences in academic performance would more strongly reflect variations in behavior.

This study contributes to the literature in three ways. First, it addresses unanswered questions about how cognitive and social skills related to academic achievement influence coordination performance. Second, it provides empirical evidence linking academic success to broader strategic abilities, such as forward-thinking and cooperation. Third, it opens new avenues for research by emphasizing the need to disentangle the specific mechanisms underlying coordination, such as problem-solving, attention, and social reasoning.

2 Methods

Population. We recruited 132 children from 6th, 7th and 8th grade in Thomas Starr King Middle School (KING), a public school in Los Angeles. The study was conducted with the University of Southern California IRB approval UP-12-00528. We distributed consent forms to parents through the school administration with an opt-out option. Students' main ethnicities at KING are Latino (55%), White (20%) and Asian (12%). Families are of very low socioeconomic status, with 75% living at or below the national poverty level. Only a minority of these students go to college. Students in the school are divided in four tracks: honors, regular, at risk, and special education. "Honors students" are high academic achievers as identified by the school. They are grouped in advanced classes whenever possible, and pooled with "regular students" otherwise. "Special education students" are individuals with special needs such as dyslexia or attention deficit. They typically have low academic performance due to their difficulty to follow the standard curriculum combined with the school's inability to give personalized education (despite their different needs, all special education students are grouped together due to budget shortages). Finally, "at risk students" are individuals with low academic performance and sometimes in-class problems of discipline. They are usually in the same class as regular students but occasionally they are pooled with special education ones. For power considerations, we grouped students in 2 tracks: 'honors' and 'regular' students are pooled in the *High* academic performance track (28, 35 and 8 students in 6th, 7th and 8th grade, respectively) whereas 'at risk' and 'special education' students are pooled in the

Low academic performance track (21, 19 and 21 students in 6th, 7th and 8th grade, respectively).

Games. Participants were anonymously paired. In half the sessions, participants were assigned a color (red or green) and played the same asymmetric battle-of-the-sexes game (**aBS**) with the same partner for 18 rounds (which, following Dal Bó and Fréchette (2018), we call a "supergame"). They then changed partners and colors and played the same 18-round supergame again. After that, they were assigned new colors (blue or yellow) and played two hawk-dove (**HD**) supergames, again with a fixed 18-round termination rule, and with change of partners and colors between the two supergames. In the other half of sessions, rules were identical except that the order was reversed (two **HD** supergames followed by two **aBS** supergames). Table 1 describes the payoff matrix of each stage game with the points used in the experiment (labels and notations will become clear later).

aBS

		green player	
		red (Y_g)	green (M_g)
red player	red (M_r)	(6,2)	(1,1)
	green (Y_r)	(1,1)	(2,4)

HD

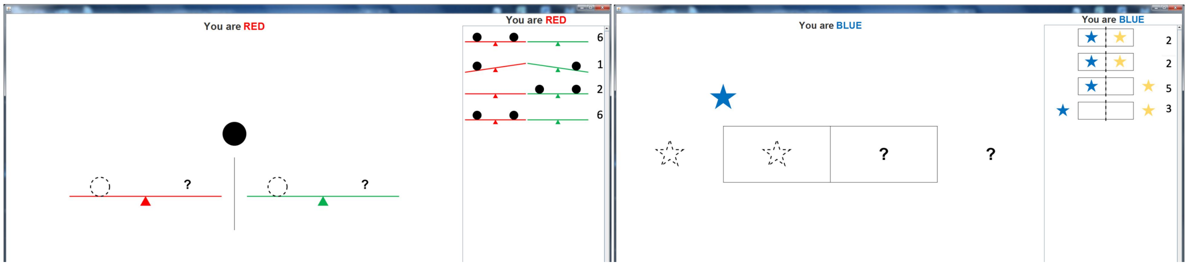
		yellow player	
		in (R_y)	out (S_y)
blue player	in (R_b)	(2,2)	(5,3)
	out (S_b)	(3,5)	(3,3)

Table 1: Payoff matrix of asymmetric battle of the sexes **aBS** and hawk-dove **HD** games

Grouping. For logistics reasons related to the availability of classrooms, teacher schedules, and students, we could not exclusively form homogeneous pairs. The experiment thus consists predominantly of pairs with two *High* (47%) or two *Low* (40%) academic achievers, but there is also a small fraction of *Mixed* pairs (13%). All pairs consist of participants in the same grade.

Presentation. It was of paramount importance to provide a simple, graphical representation of the game that all participants could easily understand, independently of their intellectual ability. This way, differences in behavior cannot be attributed to misunderstanding or confusion. Figure 1 presents screenshots of the two games with the narrative used in the experiment.

Duration and payments. The experiment consisted of two games: a short lying game analyzed in a different paper (Brocas and Carrillo, 2021b) and the present study. To avoid cross-contamination, we did not communicate the results of the first game until the end of the second. The experiment never exceeded one school period (50 minutes) including instructions and payments. Participants accumulated points that were converted into money at the rate of 3 cents per point. Since the school does not allow cash on premises, participants were paid immediately after the experiment with an Amazon e-giftcard sent to their school email. Incentives were high: average payment over the two games was $11.90, which is significantly more than typical experiments with school children. A copy of the instructions is included in Appendix A1. The data, full instructions, oTree software and its code are all available in our GitHub page `https://github.com/labelinstitute/dev_DM` under the folder "Coordination & Academic Achievement".



(a) aBS screenshot

(b) HD screenshot

Figure 1: Experimental design. **(a) aBS** "Find the Balance" game. Players have a role, red or green (here, red). They each possess a scale with their own color and a ball, which they simultaneously place on one of the scales (dotted circles for own choice and "?" for the partner's choice). If both participants place their balls on the same scale, the scale is balanced and both earn 2 points. Furthermore, the owner of the scale gets an additional 2 points (if green) or additional 4 points (if red), thus generating an asymmetry in the coordination payoff. If players place their balls on different scales, the scales are unbalanced, yielding only 1 point to each. **(b) HD** "Risky Stars" game. A blue and a yellow player (here, blue) possess a star and simultaneously decide whether to place it on or outside a common carpet (dotted stars). Placing the star outside the carpet gives 3 points for sure. Placing it on the carpet gives 2 points if the other player also places their star on the carpet and 5 points if the other player places their star outside the carpet. For both games, the right column populates the past choices of both players. It allows participants to easily track the history of outcomes and their payoffs in the supergame (here, the first four rounds).

Theory. Repeated coordination games with multiple equilibria of the stage game are particularly interesting to study empirically because theory has very limited predictive power. Indeed, according to the limit perfect "folk theorem", any feasible and individually rational payoff vector of the stage game is achievable in the finitely repeated game as the finite time horizon gets sufficiently large (Benoit and Krishna, 1985).[1] At the same time, some strategies are more natural and intuitive, and may lead to higher payoffs than others.

For the **aBS** stage-game, we call M_i the choice by player $i \in \{r, g\}$ (red or green) of 'My' favorite action (red scale for red player and green scale for green player) and Y_i the choice of 'Your' favorite action (green scale for red player and red scale for green player). For the **HD** stage-game, we call R_i and S_i the choice by player $i \in \{b, y\}$ (blue or yellow) of the 'Risky' and 'Safe' actions, respectively (see Table 1). Both are two-player coordination games with two pure strategy Nash equilibria–(M_r, Y_g) and (Y_r, M_g) in **aBS** and (R_b, S_y) and (S_b, R_y) in **HD**– and one mixed-strategy Nash equilibrium. Each player prefers a different equilibrium, and their payoff in the mixed strategy one is strictly (**aBS**) or weakly (**HD**) lower than in either of the pure strategy.

aBS captures situations where players have a common goal (a research joint venture, a new legislation) but conflicting preferences on how to achieve it (who invests the most, which provision to compromise). **HD** captures situations where countries at war, competing firms and parties in a negotiation can adopt aggressive or accomodating strategies, with the former yielding highest

[1]This contrasts with stage games that have a unique equilibrium (such as the prisoner's dilemma) where indefinite repetition is necessary to apply the folk theorem.

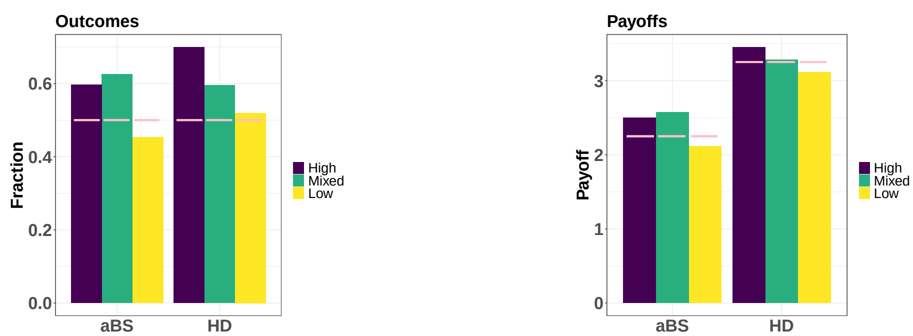
payoffs only if the rival adopts the latter.

For the repeated game version, we call *Simple Efficient Outcome (SEO)* the strict alternation between the two equilibria of the stage game: (M_r, Y_g), (Y_r, M_g), (M_r, Y_g), ... in **aBS** and (R_b, S_y), (S_b, R_y), (R_b, S_y), ... in **HD**. The idea can be summarized as "let's choose first your favorite outcome and then mine". Achieving tacit coordination has added difficulties in these games, compared to the more commonly studied symmetric battle of the sexes and stag-hunt games (Brocas and Carrillo, 2023). Indeed, the red and green players in **aBS** obtain different average payoffs under pure alternation (4 and 3, respectively). This is likely to be perceived as unfair, and make the outcome non-focal. Similarly, by choosing always S_i, a player in **HD** can avoid risks and secure a payoff of 3 that does not depend on the willingness or ability of the rival to cooperate.[2]

3 Results

3.1 Aggregate outcomes and payoffs

To study the performance of our participants, we present in Figure 2 the proportion of *equilibrium outcomes* (likelihood of playing one of the pure strategy Nash Equilibria) and the resulting average *payoffs* for each track and game, combining both supergames.[3] The first measure identifies whether individuals succeed in dynamically coordinating their actions. The second measure allows us to quantify the monetary cost of miscoordinations. In what follows, and unless otherwise stated, we perform simple two-sided t-tests for mean comparisons.



(a) Fraction of equilibrium outcomes

(b) Average payoffs

Figure 2: **(a)** Proportion of equilibrium outcomes in **aBS** ((M_r, Y_g) and (Y_r, M_g)) and in **HD** ((R_b, S_y) and (S_b, R_y)) for each track (*High* achievers, *Mixed* pairs and *Low* achievers). **(b)** Average per-round payoff in **aBS** and **HD** for each track. In both figures, behavior of a random player is represented in pink.

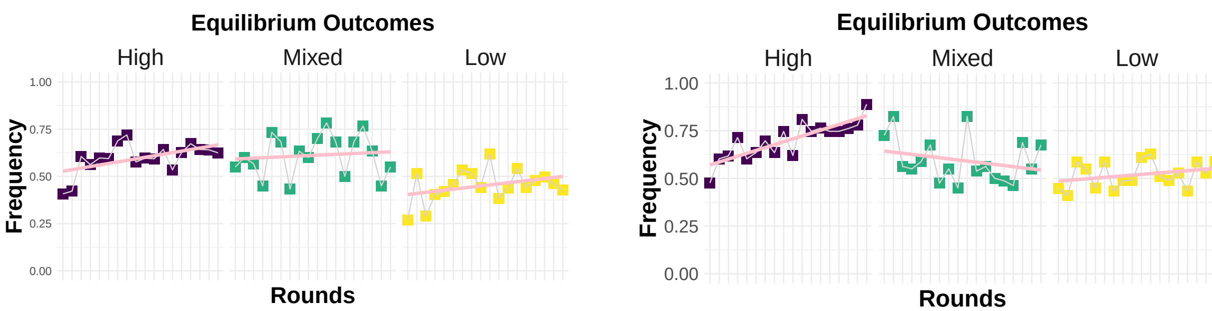
[2]By construction, any outcome where individuals play one of the pure strategy Nash equilibria in every stage is efficient and yields average payoffs in the Pareto frontier. SEO is only a natural outcome on which participants are likely to coordinate. If players wanted to reach the efficient *and fair* outcome in **aBS** (same average payoff for both players), they would need to play (Y_r, M_g) twice as often as (M_r, Y_g).

[3]As the regression analysis will demonstrate, there are no significant differences between both supergames so we combine them to increase statistical power.

These games prove challenging for our middle school participants. Under random behavior, we would observe 50% equilibrium outcomes with average payoffs of 2.25 in **aBS** and 3.25 in **HD** (pink horizontal lines). Conversely, perfect coordination on SEO would result in 100% equilibrium outcomes, with average payoffs of 3.5 and 4, respectively. Participants in the *High* track perform significantly better than random in both games, with equilibrium probabilities of 0.60 in **aBS** and 0.70 in **HD** ($p < 0.01$), though far from SEO. Their payoff is very significantly higher than random in **aBS**, but only by a small margin (2.50 vs. 2.25, $p = 0.008$). Meanwhile, the *Low* track is not statistically different from random performance, in terms of either outcomes or payoffs.[4]

We generally observe highly significant differences between *High* and *Low* tracks in both equilibrium outcomes (0.60 vs. 0.45 in **aBS**, $p = 0.006$, and 0.70 vs. 0.52 in **HD**, $p = 0.001$) and payoffs (2.50 vs. 2.11 in **aBS**, $p = 0.003$, and 3.45 vs. 3.11 in **HD**, $p = 0.002$). This provides our first indication that academic achievement is a strong predictor of performance in coordination games.

We next examine the outcome dynamics. Figure 3 illustrates the evolution from round 1 to round 18, in the proportion of groups that successfully coordinate on an equilibrium, again combining data from both supergames.



(a) aBS **(b) HD**

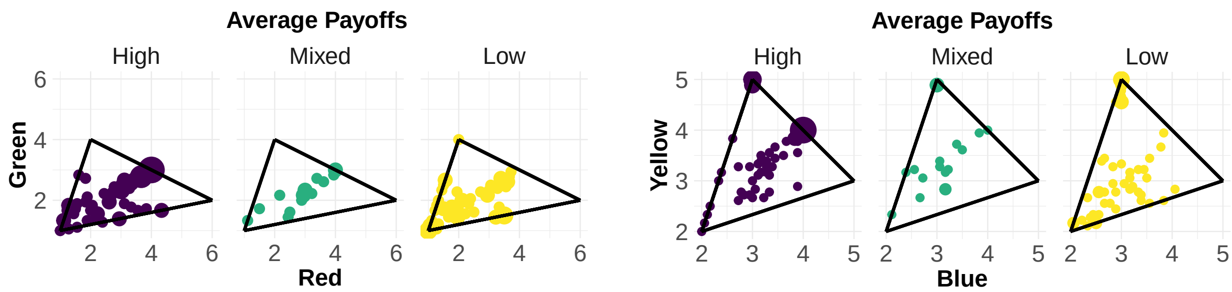
Figure 3: Evolution in average proportion of equilibrium outcomes from round 1 to round 18 of a supergame in **(a) aBS** and **(b) HD** separated by track. The best fit trend is represented in pink.

Consistent with Figure 2a, equilibrium play is more prevalent in the *High* track than in the *Low* track, with the *Mixed* track falling somewhere in between. Additionally, the slope β of the coordination trend across rounds within a supergame is positive and significant in the *High* track ($\hat{\beta}^1_H = 0.008$, $p = 0.019$ in **aBS** and $\hat{\beta}^2_H = 0.015$, $p < 0.001$ in **HD**). This corresponds to an estimated increase in coordination of 14.4% in **aBS** and 27.0% in **HD** between the first and last round of a supergame. In contrast, the *Low* track exhibits greater variability and no significant trends ($\hat{\beta}^1_L = 0.006$, $p = 0.15$ in **aBS** and $\hat{\beta}^2_L = 0.004$, $p = 0.21$ in **HD**). To compare the slopes between *High* and *Low* in each game, we combined the datasets and ran a regression with an interaction term between the round number and a dummy variable indicating the track.

[4]The behavior of the *Mixed* track looks comparable to *High* but statistical inferences are difficult to make due to the low number of observations (16 or 18 depending on the game).

The significance of the interaction term tests whether the difference in coefficients is statistically significant. We found that it was different at p $<$ 0.001 in both **aBS** and **HD**. Finally, inferences in the *Mixed* track are challenging due to the limited number of observations.

Figure 4 extends the analysis of average payoffs by examining the earnings of each pair of players in the game.



(a) aBS **(b) HD**

Figure 4: Average payoffs over all rounds of a supergame in **(a) aBS** and **(b) HD** by track. Dots correspond to pairs of players. Size of a dot is proportional to number of pairs with that payoff combination. The area inside the black lines represents the set of achievable expected payoffs in the game, with the rightmost segment (from (2,4) to (6,2) in **aBS** and from (3,5) to (5,3) in **HD**) marking the Pareto frontier.

Again consistent with Figure 2b, there is a significant difference in payoffs between the *High* and *Low* tracks. Let's define MIN as the vector of payoffs that can be easily secured by the players: (2,2) in **aBS** (achieved by each player coordinating on the rival's preferred outcome) and (3,3) in **HD** (achieved by consistently choosing the safe action). The proportion of outcomes where both players in the pair fall below MIN is 0.22 in **aBS** and 0.17 in **HD** for the *High* track, compared to 0.38 in **aBS** and 0.35 in **HD** for the *Low* track. These proportions are marginally statistically different (test of comparison of proportions, p = 0.08 in **aBS** and p = 0.05 in **HD**). Conversely, let's define MAX as the vector representing 90% of the payoff under SEO: (3.7, 2.8) in **aBS** and (3.8, 3.8) in **HD**. The proportion of outcomes where both players in the pair weakly exceed MAX is 0.17 in **aBS** and 0.29 in **HD** for the *High* track, compared to only 0.02 in **aBS** and 0.02 in **HD** for the *Low* track. These proportions are significantly different (test of comparison of proportions, p = 0.02 in **aBS** and p $<$ 0.001 in **HD**). Overall, while there is large heterogeneity in behavior within each track, the payoff gap between the *High* and *Low* tracks is considerable.

Notice that no pair of players in **aBS** achieves or approaches the efficient and fair outcome, as described in footnote 2 $(10/3, 10/3)$. Achieving such coordination likely demands both very high sophistication and a strong belief that the partner is also very sophisticated. By contrast, 13% of pairs in **HD** coordinate on or near the equilibrium preferred by one player. Such acceptance of an inequitable outcome is unusual in the literature.

Last, in Table 2, we conduct OLS regressions to analyze the determinants of performance.

	Eq. outcomes		Total payoffs	
	aBS	**HD**	**aBS**	**HD**
(const.)	10.1***	11.5***	-1.60	42.66**
	(1.00)	(1.05)	(13.89)	(14.63)
Mixed	2.39°	0.54	8.35**	1.84
	(1.38)	(1.38)	(2.93)	(3.40)
High	1.64°	2.30*	6.70**	4.54°
	(0.93)	(0.97)	(2.12)	(2.38)
HD1st	-2.90**	-2.80**	-5.69**	-5.15**
	(0.88)	(0.91)	(1.96)	(1.97)
Second	0.28	-0.21	0.40	-0.88
	(0.83)	(0.86)	(1.63)	(1.09)
Red	—	—	9.43***	—
			(1.61)	
Age	—	—	0.25**	0.12
			(0.09)	(0.10)
Male	—	—	1.36	-2.73
			(1.88)	(2.08)
Siblings	—	—	-1.23	0.24
			(2.53)	(4.23)
Adj. R^2	0.12	0.12	0.21	0.09
Num. obs.	132	132	262	262

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 2: OLS regressions to determine the effect of track (*Low* is the reference), order of play (*HD1st* is a dummy for participants who start with the **HD** game), supergame (*Second* is a dummy for the second time they play the same supergame), role (*Red* is a dummy for red players in **aBS**) and demographics (age, gender, siblings) on the outcome and payoff of players in **aBS** and **HD**.

As previously documented, *High* achievers (and to a lesser extent, *Mixed* pairs) tend to play at equilibrium more frequently and obtain higher payoffs than *Low* achievers, the reference group in the regression. There is no significant behavioral difference between the first and second supergame of either game (variable *Second*), validating our decision to pool the supergames for the aggregate analysis. Also, given the payoff asymmetry in the battle of the sexes and the lack of coordination on fair equilibria, players in the red role earn significantly higher payoffs than those in the green role (variable *Red*). Age shows a significant effect only in **aBS**, which could be due to the narrow age range in our sample. Gender and having one or more siblings do not appear to impact performance.

The most puzzling finding is the negative effect of starting with **HD** on performance and earnings in both games. While the reason for this outcome is not immediately clear, it could be that starting with **HD** leads to increased complexity or strategic confusion that carries over into subsequent gameplay.

3.2 Individual analysis

We next analyze behavior at the individual level. We consider a very extensive range of strategies that players could plausibly employ, and retain those that are empirically most frequent. A detailed explanation of our classification method can be found in Appendix A2.

The strategies most commonly used by our participants are the same across both games and are among the simplest we considered. (i) Tit-for-tat (TFT): replicate the action of the partner in the previous round; (ii) Alternate (ALT): alternate between the two actions; and (iii) Favorite (FAV): always choose the action that yields highest payoff provided the partner accommodates (M_i in **aBS**, R_i in **HD**). After categorizing players' strategies separately for the first and second supergames to account for potential behavioral changes across games, we show that, with only these three strategies, we classify 74% of decisions in **aBS** and 66% in **HD**. Each additional strategy contributes only marginally to the classification, so we do not include any other. Instead, we group the remaining 26% and 34% of players in the OTHER category.

Identifying a player's strategy within a single supergame is challenging, as we lack the counterfactual of what the player would have done if the partner had played differently. Fortunately, the overlapping of strategies is, in that respect, highly revealing. Indeed, an individual classified as both ALT and TFT (ALT&TFT) is someone who alternates between actions but also plays the action opposite to their rival in every round. Such a person achieves the SEO. In contrast, someone who is only ALT or only TFT is strategic but fails to reach the SEO. On the other hand, an individual classified as both FAV and TFT (FAV&TFT) indicates that both they and their partner consistently choose their favorite action, which leads to miscoordination and low payoffs. By contrast, someone classified only as FAV might be coordinating sometimes (or often) on their favorite outcome.

Table 3 displays the number of participants classified under each strategy (including the overlapping ones) together with their average fraction of equilibrium behavior and their average payoff.

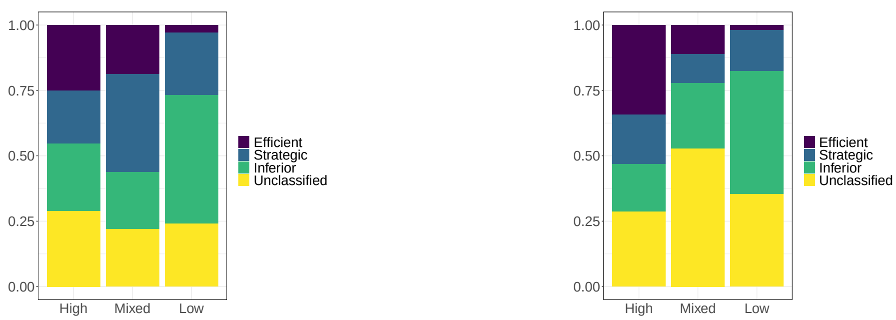
		ALT&TFT	TFT	ALT	FAV	FAV&TFT	OTHER		
		Efficient	*Strategic*		*Inferior*		*Unclassified*	*SEO*	*random*
aBS	# obs.	41	42	21	64	27	69	–	–
	% eq.	0.92	0.70	0.73	0.36	0.12	0.51	1.0	0.50
	ave. payoff	3.26	2.74	2.72	2.16	1.22	2.11	3.5	2.25
HD	# obs.	49	21	23	60	20	91		
	% eq.	0.96	0.64	0.66	0.47	0.34	0.57	1.0	0.50
	ave. payoff	3.93	3.44	3.48	2.82	2.68	3.31	4	3.25

Table 3: Number of individuals, average proportion of equilibrium choices and average payoffs under each strategy (including overlapping ones) in **aBS** and **HD**. Equilibrium and payoffs under the SEO and random choice are provided for reference.

The distribution of behavior is similar across games. Participants employ either strategies that are forward-looking and conducive to high average payoffs (*Efficient* and *Strategic*) or strategies that are mostly myopic and conducive to low average payoffs (*Inferior* and *Unclassified*).

Of the 264 observations in each game, around 15%-20% fall under ALT&TFT, an *Efficient* strategy that yields a very high fraction of equilibrium outcomes and an expected payoff close to SEO. Approximately 20%-25% of choices are classified as either ALT or TFT, which are *Strategic* choices. These strategies result in high payoffs, but below SEO due to systematic deviations by players. Then, there are 30%-35% of self-centered choices, classified as either FAV or FAV&TFT. These strategies rarely achieve coordination (especially the latter one). They lead to *Inferior* outcomes, with payoffs below the random benchmark. For the remaining 25%-35% of choices, a clear strategy is hard to discern, so we categorize them as *Unclassified*. These individuals obtain payoffs around the random outcome. They incur notable losses, though not as severe as those employing self-centered strategies.

Figure 5 presents the distribution of these four types of strategies by game and track.



(a) aBS **(b) HD**

Figure 5: Distribution of individual strategies, *Efficient* (ALT&TFT), *Strategic* (ALT or TFT), *Inferior* (FAV or FAV&TFT) and *Unclassified* (OTHER) by game (**aBS** and **HD**) and track (*High*, *Mixed*, *Low*).

Consistent with previous findings, we observe a sharp overrepresentation of *Efficient* choices among *High* achievers and *Inferior* choices among *Low* achievers in both games. Strategies are more challenging to classify in the *Mixed* track, and in **HD** compared to **aBS**.

Overall, performance is highly heterogeneous, sometimes suboptimal, but typically better in the *High* and *Mixed* tracks than in the *Low* track, with choices closer to equilibrium and higher payoffs. We performed Probit regressions to study in more detail the determinants of individual strategy choices. The results are presented in Appendix A3, and they support the findings of Figure 5. *High* achievers and *Mixed* groups are significantly more likely to adopt the *Efficient* strategy and significantly less likely to choose the *Inferior* strategy compared to *Low* achievers.

4 Discussion

This study contributes to the literature on coordination games by examining how academic achievement influences strategic behavior and coordination among middle schoolers in complex repeated-game settings, such as the asymmetric Battle of the Sexes (**aBS**) and the Hawk-Dove (**HD**) games.

Academic achievement and coordination. A central finding of the study is the strong positive link between academic achievement and performance. On average, *High* achievers outperform *Low* achievers. They are more likely to obtain superior outcomes from the outset and they also improve their coordination over the course of the supergame. This implies that academic achievement may reflect not only domain-specific knowledge but also broader cognitive skills, such as problem-solving, anticipation, and the ability to model others' behaviors—skills that are crucial not only for successful coordination in the game, but also for navigating through life. Conversely, *Low* achievers struggle with coordination which results in significant payoff losses. They do not exhibit significant trends over the course of the supergame. This gap in performance between *High* and *Low* achievers aligns with previous research suggesting that individuals with stronger cognitive skills tend to perform better in tasks requiring forward-thinking and strategic interaction (Proto et al., 2019, 2022; Fe et al., 2022). A major strength of this study is its ecological validity in evaluating cognitive skills, as it avoids reliance on abstract tasks testing narrow abilities. Instead, it uses a single measure—academic achievement—that reflects the cumulative effect of skills developed through long-term learning. However, we recognize that this choice also presents a limitation: academic achievement is a composite measure that aggregates multiple underlying skills, such as problem-solving, anticipation, and social reasoning. This breadth makes it challenging to isolate the specific mechanisms responsible for the observed differences in coordination ability. Despite this limitation, our findings emphasize the broader relevance of academic success in understanding strategic coordination and underscore its value as an ecologically valid predictor of performance in complex environments.

Strategy selection and behavior. The most successful participants use *simple yet effective* strategies that allow for efficient coordination: tit-for-tat and alternation of actions. These strategies are particularly prevalent among *High* achievers, who demonstrate a stronger ability to adopt them. In contrast, the prevalence of self-centered strategies, such as consistently choosing one's favorite action, is much higher among *Low* achievers, resulting in miscoordination and low payoffs. Relying on these uncooperative strategies suggests a difficulty to adapt to the strategic environment, further exacerbating the performance gap. It is also consistent with the existing literature which shows that poor performers in strategic games are typically individuals who focus primarily on salient, self-centered information (Costa-Gomes et al., 2001; Johnson et al., 2002; Devetag et al., 2016). Future research should explore the link between academic performance and limited attention to better understand the causes of the observed performance gap.

Heterogeneity. A second key finding of the study is the substantial heterogeneity within tracks, thus providing a nuanced view of how academic achievement influences strategic behavior. While *High* achievers generally outperform *Low* achievers, not all consistently play optimally, and some *Low* achievers occasionally demonstrate sophisticated thinking abilities that lead to efficient outcomes. This highlights the complexity of decision-making, where factors beyond academic performance—such as experience, risk-attitudes, social preferences and theory of mind—play a role.

This heterogeneity has been observed in young children and has been shown to persist over time in other strategic contexts (Brocas and Carrillo, 2021a). *High* achievers may benefit from more challenging tasks that deepen strategic thinking, while *Low* achievers might need foundational support in trust-building and cooperation. Tailored interventions and opportunities could help improve outcomes for those who struggle with coordination tasks.

Performance in mixed pairs. An intriguing aspect of the study is the performance of the Mixed group. These pairs show intermediate performance (closer to *High* than to *Low* achievers), suggesting that *Low* achievers benefit from interacting with *High* achievers, who likely provide guidance that improves coordination. This aligns with research on peer learning, where exposure to more advanced peers promotes cognitive development (Hmelo-Silver, 2013). Although it is based on a small sample, this finding suggests that pairing *Low* achievers with *High* achievers can enhance performance in educational settings. Further research is needed to explore the dynamics of these interactions and optimize their benefits.

Inequity. One surprising finding is the prevalence of inequitable outcomes by participants, particularly in **HD**, where around 13% consistently coordinate on outcomes favoring their partner at their own expense but also in **aBS**, where coordination on efficient and fair outcomes is non-existent. This contrasts with the common expectation of inequity aversion, with people typically resisting unequal payoffs, even at personal cost (Fehr and Schmidt, 1999). Several factors could explain this behavior. Younger individuals may be less sensitive to inequity or prioritize maintaining cooperation over fairness. The dynamics of **HD**, where one strategy involves a smaller but safe payoff, might also lead players to view these outcomes as acceptable, even if inequitable. An alternative explanation could be that our participants are inequity averse but find it difficult to identify a way to achieve outcomes that are both efficient and fair. Further research could explore how factors like communication, stakes or environmental complexity influence inequity tolerance and whether this behavior persists into adulthood.

Order of games and behavior. Starting with the Hawk-dove game negatively impacts performance in both games. One possible explanation is psychological priming: **HD** involves conflict, which may prime participants for competitive behavior. Risk-aversion may also play a role, as **HD** encourages safe, cautious strategies that do not translate well into **aBS**. Finally, emotional responses from the conflict in **HD**—such as frustration or distrust—could linger, impairing trust and cooperation in subsequent games. In summary, the observed order effect likely arises from a mix of cognitive, emotional, and strategic factors. Further research could explore how these elements interact to affect performance.

Education. A natural question would be to investigate how interventions aimed at improving strategic thinking and coordination skills might mitigate the documented performance gap. Understanding whether targeted training can improve outcomes of *Low* achievers could have important implications for educational practices. Also, *Mixed* groups generally outperform *Low* groups. This suggests that offering opportunities for collaborative learning and problem-solving may be a way for educators to help support students who struggle with strategic thinking.

Limitations. While this study provides valuable insights into the relationship between academic achievement and coordination performance, there are several limitations worth noting. First, the study is restricted to middle school students, which may limit the generalizability of the findings to other age groups. Additionally, the study does not explore the specific cognitive or social factors that drive the performance differences between *High* and *Low* achievers. Future research could benefit from examining the role of specific cognitive abilities, such as working memory, cognitive flexibility, or theory of mind, in shaping coordination behavior to provide a clearer understanding of the processes at play.

References

Masaki Aoyagi and Guillaume Fréchette. Collusion as public monitoring becomes noisy: Experimental evidence. *J. of Economic theory*, 144(3):1135–1165, 2009.

Jean-Pierre Benoit and Vijay Krishna. Finitely repeated games. *Econometrica*, 53(4):905–22, 1985.

Isabelle Brocas and Juan D Carrillo. Steps of reasoning in children and adolescents. *Journal of Political Economy*, 129(7):2067–2111, 2021a.

Isabelle Brocas and Juan D Carrillo. Self-serving, altruistic and spiteful lying in the schoolyard. *Journal of Economic Behavior & Organization*, 187:159–175, 2021b.

Isabelle Brocas and Juan D Carrillo. Dynamic coordination in efficient and fair outcomes: a developmental perspective. *Available at SSRN 4467556*, 2023.

Gabriele Camera, Marco Casari, and Maria Bigoni. Cooperative strategies in anonymous economies: An experiment. *Games and Economic Behavior*, 75(2):570–586, 2012.

Colin Camerer. *Behavioral game theory: Experiments in strategic interaction.* Princeton University Press, 2003.

David J Cooper and Roberto A Weber. Recent advances in experimental coordination games. *Handbook of experimental game theory*, pages 149–183, 2020.

Miguel Costa-Gomes, Vincent P Crawford, and Bruno Broseta. Cognition and behavior in normal-form games: An experimental study. *Econometrica*, 69(5):1193–1235, 2001.

Pedro Dal Bó and Guillaume R Fréchette. On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, 56(1):60–114, 2018.

Giovanna Devetag, Sibilla Di Guida, and Luca Polonio. An eye-tracking study of feature-based choice in one-shot games. *Experimental Economics*, 19:177–201, 2016.

Catherine C Eckel and Rick K Wilson. Social learning in coordination games: does status matter? *Experimental Economics*, 10:317–329, 2007.

Eduardo Fe, David Gill, and Victoria Prowse. Cognitive skills, strategic sophistication, and life outcomes. *Journal of Political Economy*, 130(10):2643–2704, 2022.

Ernst Fehr and Klaus M Schmidt. A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868, 1999.

Cindy E Hmelo-Silver. The international handbook of collaborative learning. 2013.

Christos A Ioannou and Julian Romero. A generalized approach to belief learning in repeated games. *Games and Economic Behavior*, 87:178–203, 2014.

Matthew O Jackson and Yiqing Xing. Culture-dependent strategies in coordination games. *Proceedings of the National Academy of Sciences*, 111(supplement_3):10889–10896, 2014.

Eric J Johnson, Colin Camerer, Sankar Sen, and Talia Rymon. Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of economic theory*, 104(1):16–47, 2002.

Marco Lambrecht, Eugenio Proto, Aldo Rustichini, and Andis Sofianos. Intelligence disclosure and cooperation in repeated interactions. *American Economic Journal: Microeconomics*, 16(3): 199–231, 2024.

Richard D McKelvey and Thomas R Palfrey. Playing in the dark: Information, learning, and coordination in repeated games. *Caltech*, 2001.

Joanne Peryman and David Kelsey. Ambiguity when playing coordination games across cultures. *Theory and Decision*, 90(3):485–505, 2021.

Eugenio Proto, Aldo Rustichini, and Andis Sofianos. Intelligence, personality, and gains from cooperation in repeated interactions. *Journal of Political Economy*, 127(3):1351–1390, 2019.

Eugenio Proto, Aldo Rustichini, and Andis Sofianos. Intelligence, errors, and cooperation in repeated interactions. *The Review of Economic Studies*, 89(5):2723–2767, 2022.

Julian Romero and Yaroslav Rosokha. The evolution of cooperation: The role of costly strategy adjustments. *American Economic Journal: Microeconomics*, 11(1):299–328, 2019.

SI. Appendix

Appendix A1. Sample of instructions

Hi, my name is ________, and these are my helpers [Introduce helpers]. Today, we are going to play a few games with you. In all the games, you will earn points that will be placed in your virtual wallet. At the end of the experiment you will be paid 3 cents for each point you obtained with an Amazon gift card. You will get several hundred points, so you will be able to get a nice gift card. Are you excited?

Risky stars

You are going to play a game called "Risky stars" with one other person in the room. The computer will decide with whom you play that game. One of you will be BLUE and the other will be YELLOW. The computer also decides who is BLUE and who is YELLOW. If you are BLUE, your screen looks like this. A visualization of all the slides is provided in Figure 6

[SLIDE 1]

At the top of the screen, it says you are BLUE. In the middle of the screen there is a carpet divided in two. You own the BLUE star. You need to decide whether to put your star on the carpet or outside the carpet. If you are YELLOW, your screen looks like this.

[SLIDE 2]

It says you are YELLOW at the top and you see the same carpet as your partner. You own the YELLOW star and you need to decide whether to put your star on the carpet or outside the carpet. Here you can see the screen of BLUE and the screen of YELLOW.

[SLIDE 3]. Now, how do you get points?

[SLIDE 4]

If you put your own star outside the carpet, you get 3 points, no matter what your partner does. If you put it on the carpet, what you get depends on what your partner does. If he also puts his star on the carpet, neither of you fits well and you both earn 2 points. But if he puts it outside the carpet, then you have all the space for yourself and you earn 5 points (while he earns 3 points for being outside the carpet). This information will remain here [point] during your choices.

Now, this is very important. You are going to play many times with the same partner. Each time, you will make your choices at the same time. This means that you will not know what your partner did when you make your choice. It is only after both of you have made a choice, that you will both know what each of you did and how many points you got. This will appear in the right column of your screen.

[SLIDE 5]

For instance, this is what your screen may look like after 4 rounds. [explain what happened in the first, second, and third round ...]. You have accumulated a total of 12 points so far, but you are going to play many more rounds.

OK. The computer will now select partners. The points you get depending on where the stars are placed will remain visible on the big screen [go back to SLIDE 11]. When you are ready, make your choices. Just tap on the screen where you want to place your star.

[After supergame 1] The game has ended. You can see on your screen the points accumulated. The computer will now select new partners and you will play the same game again. You may have the same role as before but note that it does not matter since the points are the same for both colors.

[After supergame 2] The game has ended. You can see on your screen the points accumulated.

Find the balance

You are going to play with one other person in the room a game called "Find the balance". As before, the computer will decide with whom you play this game. One of you will be RED and the other will be GREEN. The computer decides who is RED and who is GREEN. If you are RED, your screen looks like this.

[SLIDE 6]

At the top of the screen, it says you are RED. You own the RED scale and the black ball in the middle of your screen. You can also see a GREEN scale, which belongs to your partner. You need to decide whether to put your ball on the dotted circle of your RED scale or on the dotted circle of your partner's GREEN scale. If you are GREEN, your screen looks like this.

[SLIDE 7]

It says you are GREEN at the top. You own the GREEN scale and the black ball in the middle of your screen. You can also see a RED scale which belongs to your partner. You need to decide whether to put your ball on the dotted circle of your GREEN scale or on the dotted circle of your partner's RED scale. Here you can see the screen of RED and the screen of GREEN.

[SLIDE 8]. Now, how do you get points?

[SLIDE 9]

If both balls are put on the RED scale, the scale is balanced. Player RED gets 6 points and player GREEN gets 2 points. If both balls are put on the GREEN scale, then the scale is again balanced. Player RED gets 2 points and player GREEN gets 4 points. If the balls are put on different scales, the scales are unbalanced. The balls fall and each player gets only 1 point. This information will remain here [point] during your choices.

Notice that RED gets more points when both balls are on the RED scale than the points GREEN gets when both balls are on the GREEN scale. Now this is very important. You will play many rounds with the same partner. In each round, you will make your choices at the same time. This means that you will not know what your partner did when you make your choice. However, after both of you have made a choice, you will both know what each of you did and how many points you got. This will appear in the right column of your screen.

[SLIDE 10]

For instance, this is what your screen may look like after 4 rounds. [explain what happened in the first, second, and third round ...]. You have accumulated a total of 15 points so far, but you are going to play many more rounds. All right, are you ready to play? The computer will now select partners. The points you get depending on where the balls are placed will remain visible on the big screen [go back to SLIDE 16]. When you are ready, make your choices. Just tap on the screen where you want to place your ball.

[At the end of the 1st game] The game has ended. You can see on your screen the points you have accumulated. The computer will now select new partners and you will play the same game again. If you were RED before, you will be GREEN and if you were GREEN you will be RED. Remember the game is the same, but your color role is different and your partner is different.

[At the end of the 2nd game] The game has ended. Please answer a few questions and we are done.

Payments. We will now call you one by one and tell you how much money you earned. You can tell your friends how much you got or not. It is totally up to you. You will get today an email from Amazon with an amazon e-giftcard for that amount. If you don't receive it, please let us know. Thanks for playing with us!

Figure 6: Slides to accompany instructions

Appendix A2. Methodology to determine individual strategies

We investigate individual heterogeneity by classifying our participants according to their strategy in each supergame. We employ the following methodology. For each supergame, we study the 15 choices in rounds 4 to 18. We assume that choices in rounds 1 and 2 are explorations and therefore do not consider them. We use the outcome in round 3 as the anchor (initial condition) for the strategy, and consider a large number of possible strategies. We assign to each player the strategy for which the number of deviations in rounds 4 to 18 is smallest, as long as it is no greater than 3. If the number of deviations is the same for two or more strategies, the subject is classified at the intersection. Players with more than 3 deviations from all the strategies remain unclassified. There are only minor variations in results if we use rounds 2 or 4 as anchor and/or if we allow 2 or 4 deviations. In the analysis, the subject's initial choice is left unspecified for some strategies. This is not theoretically rigorous but is consistent with using round 3 as the anchor of the empirical analysis.

As mentioned above, the behavior of a player in a supergame may be compatible with several strategies. With a large number of supergames, one could disentangle between strategies by studying the behavior against different partners, adapting some of the sophisticated techniques developed in the repeated prisoner's dilemma literature.[5] Such analysis is not possible with only two supergames. Our best approach is to study overlaps of strategies as we discuss in the text. Also, we allow for different strategies across supergames because we expect some participants to change their behavior between the first and second time they play.

Strategies in aBS

SEO in **aBS** consists in alternating between the two Nash Equilibria: (M_r^t, Y_g^t), (Y_r^{t+1}, M_g^{t+1}), (M_r^{t+2}, Y_g^{t+2}), etc. (where we use the superscript t for round t). We are interested in strategies that can help sustain SEO, but also in strategies that may be chosen intuitively even though they do not result in SEO. Table 4 reports some possible strategies from simplest to most sophisticated.

strategy	description
(1) FAV	play always M_i^t
(2) YOUR	play always Y_i^t
(3) ALT	alternate between M_i^t and Y_i^t
(4) TFT	tit-for-tat: replicate the action of the partner in the previous round
(5) TRIG	grim-trigger: play the action consistent with SEO if all past outcomes are consistent with SEO and play M_i^t forever otherwise
(6) 2TFT	two-tit-for-tat: alternate between playing once the red equilibrium and twice the green one
(7) REV	reverse tit-for-tat: reverse the choice of the partner in the previous round
(8) FORG	forgiving trigger: play M_i^t unless the last round outcome was (M_i^{t-1}, Y_j^{t-1})
(9) TEACH	play Y_i^t unless the last round outcome was (Y_i^{t-1}, M_j^{t-1})
(10) TEST	play M_i^t unless the last round outcome was (Y_i^{t-1}, M_j^{t-1})

Table 4: Some simple strategies in **aBS**

Strategies (1)-(2)-(3) can be played by naïve players with little understanding of the partners' incentives as well as by strategic players whose objective is to reach an equilibrium (insisting on the best possible for themselves, agreeing on the best for the partner, or targeting the SEO, respectively). Strategies (4)-(5) are typical in other games and may result in SEO but also collapse into (M_r, M_g) depending on the

[5]For example, Camera et al. (2012) trades-off goodness of fit and number of strategies, Aoyagi and Fréchette (2009) conduct Maximum Likelihood Estimation of best fitting strategies and Romero and Rosokha (2019) perform a direct elicitation of strategies.

partner's behavior. Strategy (6) is special to this asymmetric game since it could result in the efficient and fair equilibrium, where both players obtain the same expected payoff in the Pareto frontier. Strategy (7) seeks to repeatedly coordinate in the same static Nash equilibrium (either always exploiting the partner or always giving-in). The remaining strategies capture a variety of strategic behaviors: (8) is similar to (5) except that it forgives after one period, (9) attempts to teach SEO by playing Y after a deviation, and (10) is the opposite of (9) and similar to (7), in that it attempts to exploit partners but gives in to selfish ones.

Choosing how many strategies to include can be delicate and subjective. In this case, however, it was reasonably simple. Indeed, with only three strategies–FAV, ALT, and TFT–we can account for the choices of 73.9% of observations. Adding TEST (the most employed of the remaining ones) would increase the proportion of choices accounted to only 75.8% while including all 10 strategies would classify a total of 80.3%. Keeping only FAV, ALT, and TFT thus seemed a good compromise (note in particular that we do not observe a single instance of the efficient and fair strategy 2TFT).

Strategies in HD

SEO in **HD** consists also in alternating between the two Nash Equilibria: (R_b^t, S_y^t), (S_b^{t+1}, R_y^{t+1}), etc. In fact, and despite some important differences between the two games, the different types of behaviors are similar in the two cases. We therefore decided to consider the exact same potential strategies as in **aBS** (Table 4), except that we replace M_i with R_i (the action that could potentially lead to highest payoff) and Y_i with S_i (the choice that most closely yields to the opponent's favorite choice). We also did not consider strategy (6), since it had no meaning in the symmetric **HD** (it was not selected by anyone anyway).

Once again, the selection of the strategies to retain was reasonably simple. With the same three strategies as before–FAV, ALT, and TFT–we can account for the choices of 65.6% of observations. Adding TEACH would increase the proportion of choices accounted to 69.7% while including all 9 strategies would classify 72.3%. We ultimately decided to keep only FAV, ALT, and TFT because they were the same as in the other game. Besides, TEACH was not the fourth most employed strategy in **aBS** and it only added 4.1% of classified observations.

Appendix A3. Probit regressions on individual strategies

We present in Table 5 Probit regressions to study the determinants of individual strategy choices, using the same variable as in Table 2.

In support of previous results, we find that *High* achievers are significantly more likely to select the *Efficient* strategy and significantly less likely to select the *Inferior* strategy than *Low* achievers. A similar effect is found for the *Mixed* group. Age is the only demographic variable that has a (minor) effect, and only in **HD**. As in Table 2, we obtain the puzzling result that individuals who start with **HD** perform worse in both games. The only novel finding is that, contrary to previous results, there is a noticeable improvement in the likelihood of playing the *Efficient* strategy between supergames 1 and 2 in both games.

	aBS		HD	
	Prob. *Efficient*	Prob. *Inferior*	Prob. *Efficient*	Prob. *Inferior*
(const.)	-3.41	2.16	-3.08	3.60^{*}
	(2.11)	(1.67)	(2.76)	(1.65)
Mixed	1.05^{**}	-0.74^{*}	0.67	-0.67°
	(0.39)	(0.35)	(0.61)	(0.37)
High	1.31^{***}	-0.52^{*}	1.78^{***}	-0.87^{***}
	(0.35)	(0.23)	(0.45)	(0.24)
HD1st	-0.36	0.89^{***}	-2.35^{***}	0.87^{***}
	(0.28)	(0.23)	(0.42)	(0.25)
Second	0.47^{*}	-0.01	0.74^{***}	0.20
	(0.19)	(0.14)	(0.20)	(0.14)
Red	-0.14	-0.23	—	—
	(0.18)	(0.14)		
Age	0.01	-0.02	0.01	-0.03^{*}
	(0.01)	(0.01)	(0.02)	(0.01)
Male	0.29	-0.14	-0.01	0.33
	(0.25)	(0.20)	(0.29)	(0.23)
Siblings	0.20	0.08	0.18	-0.45
	(0.33)	(0.32)	(0.43)	(0.29)
Adj. R^2	207.0	303.5	152.2	268.9
Num. obs.	262	262	262	262

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 5: Probit regressions to determine the effect of track, order of play, supergame, role and demographics on the likelihood of playing *Efficient* and *Inferior* strategies in **aBS** and **HD**.