Journal of the Economic Science Association (2023) 9:16–33
https://doi.org/10.1007/s40881-022-00125-z

ORIGINAL PAPER

Do we all coordinate in the long run?

Manja Gärtner[1,2] · **Robert Östling**[3] · **Sebastian Tebbe**[4]

Received: 22 October 2020 / Revised: 7 December 2022 / Accepted: 27 December 2022 /
Published online: 6 February 2023
© The Author(s) 2023

Abstract

Players often fail to coordinate on the efficient equilibrium in laboratory weak-link coordination games. In this paper, we investigate whether such coordination failures can be mitigated by increasing the number of rounds or altering per-period stakes. We find that neither time horizon nor stakes affect equilibrium selection. In contrast to previous findings, players are not more likely to play above the previous period's minimum choice when the horizon is longer or per-period stakes lower. We also investigate which socio-demographic factors and behavioral traits correlate most strongly with play both in the first round and in subsequent rounds. Cognitive ability as measured by a cognitive reflection test stands out as the characteristic that is most strongly associated with efficient coordination.

Keywords Coordination games · Weak-link games · Coordination failure · Strategic risk

JEL Classification C72 · C92

Replication for this paper is publicly available at https://doi.org/10.17605/osf.io/n8hpj.

✉ Robert Östling
robert.ostling@hhs.se

Manja Gärtner
manja.gaertner@gmail.com

Sebastian Tebbe
sebastian.tebbe@iies.su.se

[1] Competition and Consumers Department, German Institute for Economic Research (DIW Berlin), Berlin, Germany

[2] Division of Economics, Department for Management and Engineering, Linköping University, Linköping, Sweden

[3] Department of Economics, Stockholm School of Economics, Stockholm, Sweden

[4] Institute for International Economic Studies, Stockholm University, Stockholm, Sweden

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

1 Introduction

Several scholars have claimed that inefficiencies arising from a failure to play an efficient equilibrium are ubiquitous in many economic settings. For example, coordination failure may hinder economic development (Kremer, 1993), lead to bank runs (Diamond & Dybvig, 1983) and affect business cycle dynamics (Bryant, 1983). It has also been argued that coordination failures are common within organizations (Knez & Camerer, 1994). A frequently used tool to study coordination failures in the laboratory is the weak-link coordination game. In the weak-link game, each player simultaneously chooses an effort level and the payoff is increasing in the minimum effort of the group (Van Huyck et al., 1990). Providing effort is costly, however, so players have an incentive not to exceed the group minimum. Players hence face a trade-off between exerting high effort to increase the group's payoff and avoiding unnecessarily high costly effort above the group minimum. Because the best response is to play the group minimum, the game has as many pure strategy Nash equilibria as there are effort levels.

The weak-link game is susceptible to coordination failure as the mere suspicion that one other player provides low effort is an incentive to exert low effort. In the laboratory, players typically play ten rounds and generally fail to coordinate on the efficient equilibrium (Van Huyck et al., 1990, 1991; Cooper et al., 1990, 1992; Knez & Camerer, 1994; Goeree & Holt, 2001, 2005). In sharp contrast, Berninghaus and Ehrhart (1998) found that players manage to coordinate when they interact over more rounds. In their experiment, players interact over 10, 30 or 90 rounds. When players interact over 10 rounds, only one out of six groups coordinates on the efficient equilibrium, whereas five out of six groups coordinate efficiently when they interact over 90 rounds. To the best of our knowledge, this finding has not been replicated, although related ideas have been explored in the literature. For example, inspired by prisoners' dilemma experiments in continuous time (Friedman & Oprea, 2012; Bigoni et al., 2015), Leng et al. (2018) show that players do not coordinate efficiently when playing the weak-link game in continuous time.[1]

Intuitively, a longer horizon might make players more willing to sacrifice payoffs making above-minimum effort choices during a couple of initial periods to improve coordination in the longer term. Along with this intuition, we show in Section C in the Online Appendix that, in the presence of strategic uncertainty, more game repetitions make strategies involving early above-minimum choices more attractive relative to always playing the previous period's minimum.[2] However, because

[1] Another related study is Chaudhuri et al. (2009) who study up to nine non-overlapping generations playing a weak-link game ten times in each generation. Each generation is allowed to pass on advice to the next generation. Most groups did not coordinate efficiently in their experiment except for one treatment in which the content of the advice was experimentally manipulated.

[2] An alternative motivation why the horizon could be relevant for coordination is that subjects stuck in a bad equilibrium occasionally experiment by choosing a higher effort level. If all subjects in a group happen to experiment simultaneously, which is more likely to happen the longer the game is repeated, a more efficient equilibrium would be reached. As pointed out by Crawford (1991, p. 53), however, the argument can go the other way if players experiment by playing both above and below the previous period's minimum.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Berninghaus and Ehrhart (1998) keep the total stakes constant it is unclear whether their findings are best explained by lower per-period stakes or longer horizon. Lower per-period stakes could, for example, affect coordination by increasing players willingness to experiment. In this paper, we set out to test whether Berninghaus and Ehrhart's (1998) findings are due to longer horizon or lower stakes.

Although our study was designed to be powered to detect effect sizes substantially smaller than those reported in Berninghaus and Ehrhart (1998), we find that neither the time horizon nor per-period stakes affect coordination. Our findings are not different from those in Berninghaus and Ehrhart (1998) with respect to behavior in the first period. The results rather differ in how players adjust after the initial period. In our experiment, above-minimum players tend to adjust to the minimum effort level in previous rounds irrespectively of stakes or horizon, whereas Berninghaus and Ehrhart (1998) found that minimum players are more likely to increase effort choices with a longer horizon.

We also study how demographic characteristics and behavioral traits are related to coordination in the weak-link game. Previous experimental research has primarily focused on structural determinants of coordination failures such as the payoff-risk ratio, group size, and feedback (see the comprehensive review by Devetag and Ortmann, 2007). Less is known about the role of behavioral and socio-demographic factors in coordination games.[3] The individual characteristic most robustly related to high initial period choices and final group coordination in our experiment is cognitive ability (measured by a cognitive reflection test). Our study was conducted in Copenhagen and we replicate Engelmann and Normann's (2010) finding that Danish subjects tend to make higher effort choices in the initial period, but in our experiment it does not result in higher coordination in groups with a higher fraction of Danish subjects.

2 Experimental design

We study a simultaneous-move weak-link game with $N = 4$ players. In each round, player $i \in \{1, 2, 3, 4\}$ chooses an effort level $e_i \in \{1, 2, 3, 4, 5, 6, 7\}$. Let $e := (e_1, e_2, e_3, e_4)$ denote a (pure) strategy profile, and $\underline{e} := min\{e_1, e_2, e_3, e_4\}$ the minimum effort level in a group. The payoff function of player i is

$$\pi_i(e_i, \underline{e}) = E + a\underline{e} - ce_i, \tag{1}$$

where $a > c > 0$ and E is a constant. Any strategy profile $(e, \ldots, e)$ with $e \in \{1, 2, 3, 4, 5, 6, 7\}$ is a pure-strategy Nash equilibrium. These seven equilibria can be Pareto-ranked with $e = (7, 7, 7, 7)$ being the Pareto dominant equilibrium and $e = (1, 1, 1, 1)$ is dominated by all other equilibria.

[3] For example, behavior in coordination games has been linked to risk preferences (Heinemann et al., 2009; Al-Ubaydli et al., 2013; Büyükboyaci, 2014; Neumann & Vogt, 2009), patience (Al-Ubaydli et al., 2013), cognitive ability (Jones, 2008; Proto et al., 2022; Al-Ubaydli et al., 2016) and trust (Devetag, 2009).

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Table 1 Payoff matrix with low stakes (in DKK)

	Minimum effort level $\underline{e}$						
Individual effort e_i	7	6	5	4	3	2	1
7	0.90	0.75	0.60	0.45	0.30	0.15	0.00
6		0.85	0.70	0.55	0.40	0.25	0.10
5			0.80	0.65	0.50	0.35	0.20
4				0.75	0.60	0.45	0.30
3					0.70	0.55	0.40
2						0.65	0.50
1							0.60

Let e_i^t denote player i's effort choice in period t, $\underline{e}^t$ the group minimum in period t and π_i^t the resulting payoff in period t. After each round, players receive feedback only about their own payoff π_i^t and the group minimum $\underline{e}^t$. Consequently, when making a choice in period t, each subject has information about the history of their own effort choices $\{e_i^s\}_{s=1,\ldots,t-1}$, the minimum effort level $\{\underline{e}^s\}_{s=1,\ldots,t-1}$ in the group and own payoff $\{\pi_i^s\}_{s=1,\ldots,t-1}$. This feedback information environment follows much of the experimental literature on coordination games (Van Huyck et al., 1990, 1991, 1993).

The experiment employs a 2×2 design, which varies the treatment dimensions along the time horizon (short/long) and stake size (low/high). Subjects are randomly assigned within sessions to one of the four treatments. The first treatment dimension varies the number of rounds. Subjects play 10 periods in the short horizon treatment and 50 times in the long horizon treatment. Because Berninghaus and Ehrhart (1998) found that many groups managed to coordinate efficiently already after 30 rounds, we deemed 50 rounds to be sufficient for coordination.

The second treatment dimension varies the stakes in each round. The payoff function parameters in the low stakes treatment are given by $E^{\mathrm{l}} = 0.55$, $a^{\mathrm{l}} = 0.15$ and $c^{\mathrm{l}} = 0.10$ and for the high stake game by $E^{\mathrm{h}} = 2.75$, $a^{\mathrm{h}} = 0.75$ and $c^{\mathrm{h}} = 0.50$. The payoff tables for the low and high stake treatment groups are shown in Tables 1 and 2. The total payoff is simply the sum of payoffs in all rounds. Subjects in the high stakes treatments can earn a maximum of 4.50 DKK per round ($\approx$ \$0.81), while subjects in the low stakes treatments can earn a maximum of 0.90 DKK per round ($\approx$ \$0.16).[4]

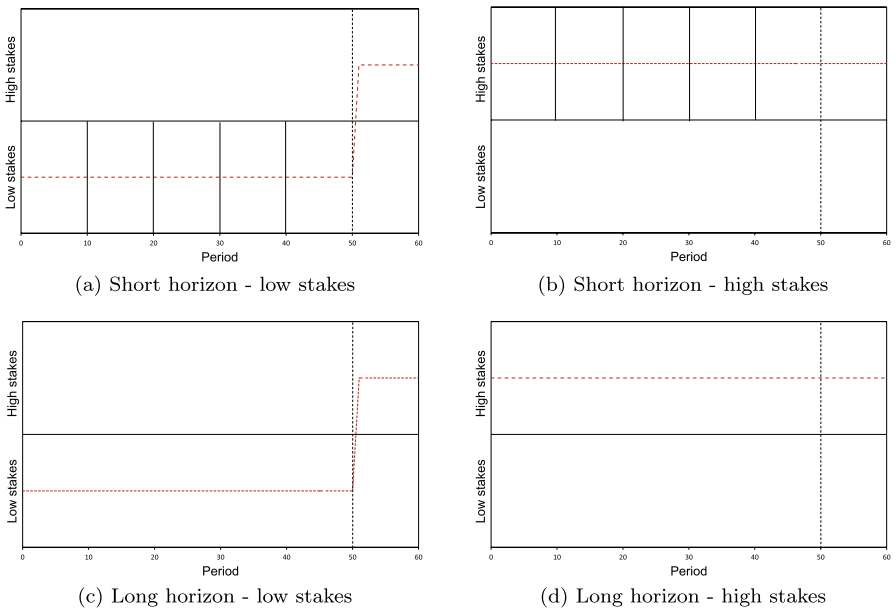
In contrast to Berninghaus and Ehrhart (1998), we vary per-period stakes independently of the horizon. Their design keeps total stakes constant so that the per-period payoff in the short horizon treatment is three times the per-period payoff in the treatment with 30 rounds and nine times larger than in the treatment with 90 rounds. Our high stakes treatment has per-period payoffs which are five times larger than our low stakes treatment, which corresponds to a five times longer horizon in

[4] We convert Danish kroner to US dollars using the exchange rate from December 31, 2013 (0.18 USD/DKK).

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Table 2 Payoff matrix with high stakes (in DKK)

	Minimum effort level $\underline{e}$						
Individual effort e_i	7	6	5	4	3	2	1
7	4.50	3.75	3.00	2.25	1.50	0.75	0.00
6		4.25	3.50	2.75	2.00	1.25	0.50
5			4.00	3.25	2.50	1.75	1.00
4				3.75	3.00	2.25	1.50
3					3.50	2.75	2.00
2						3.25	2.50
1							3.00



(a) Short horizon - low stakes

(b) Short horizon - high stakes

(c) Long horizon - low stakes

(d) Long horizon - high stakes

Fig. 1 Treatment design. The graphs present the four different treatments. The red dotted lines specify the stake size in each round, the vertical solid black lines represent the re-matching to new group constellations and the black dotted line indicate the start of unexpected last 10 rounds with high stakes and fixed groups

the long horizon treatment. This design feature allows us both to study the effect of changing the horizon while keeping per-period stakes constant, and the effect of increasing the horizon while keeping total stakes constant as in Berninghaus and Ehrhart (1998).

Figure 1 summarizes our four treatments. In the short horizon treatments, subjects first play the weak-link game with low (Fig. 1a) or high stakes (Fig. 1b) ten times with a fixed group. In the long horizon treatments, subjects first play the game 50 times with a fixed group, with either low (Fig. 1c) or high stakes (Fig. 1d). This

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

summarizes the essential part of our 2×2 experimental design and in our analyses below, unless otherwise noted, we contrast the first 50 rounds in the long horizon treatments to the first 10 rounds in the short horizon treatments. In the beginning of the experiment, subjects were informed that the experiment would consist of additional tasks after the first 10 or 50 rounds of the weak-link game, but not what those tasks would consist of. We did not inform subjects about those tasks because we did not want expectations about future tasks to influence their behavior in the weak-link game.

In the short horizon treatments, it was announced after the tenth round that subjects would play an additional 40 times and that they would be randomly re-matched with new group members every 10 periods, which is represented by the vertical solid black lines in Fig. 1a and b. After 50 periods, subjects in all treatments are informed that they will play an additional 10 periods with high stakes with the same group members as in the last round.[5]

Our main motivation for letting subjects in the short horizon treatment play an additional 40 periods is to equalize the duration of the experiment across treatments, but it also allows us to check if there is an effect of re-matching subjects. The unexpected last 10 periods are included to ensure that subjects in the low stakes treatment get a sufficiently high final payoff, but it also allows us to test whether potentially successful coordination in treatments with low stakes would persist when increasing stakes.

There are some differences between our experimental design and the most closely related studies (Berninghaus & Ehrhart, 1998; Engelmann & Normann, 2010). A detailed comparison is shown in Table A1 of the Online Appendix, but the most important difference apart from our 2×2 design is that we use smaller groups, a different benefit-cost ratio (a/c) and a different fixed-to-variable benefit ratio (E/a). Our groups consist of four members while Berninghaus and Ehrhart (1998) used groups of eight. Lowering the group size shifts the empirical distribution of coordination towards more efficient levels (Van Huyck et al., 2007) and also makes above-minimum repeated-game strategies more attractive (this is apparent from the analysis in Section C in the Online Appendix). Previous experimental literature also suggests that coordination failures are influenced by subject-pool effects. Particularly relevant in our context is that Danish subjects have shown to be much more likely to coordinate on the Pareto-efficient equilibrium even in the short run (Engelmann & Normann, 2010). Because we ran our experiment in Copenhagen in Denmark, we used a smaller variable benefit-cost (a/c) ratio of 1.5 to make coordination more challenging. We set the fixed payoff E to avoid negative payoffs. One additional difference compared to Berninghaus and Ehrhart (1998) is how the time limit in each period was enforced. Berninghaus and Ehrhart (1998) required subjects to choose within 10, 30 or 90 s and the default choice if no choice was made was the choice from the preceding period. Subjects in their experiments were also allowed to revise their decision within the time limit. In contrast, we did not enforce a time limit,

[5] In one of the sessions (session 0), the additional 10 rounds were not played.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

but subjects were encouraged to make a choice within 15 s.[6] Once a choice was made it could not be revised. We kept the per-period length constant across treatments to eliminate potential effects arising from differences in time pressure across treatments.

After subjects had finished the main part of the experiment, they participated in an incentivized Ultimatum Game and were given the opportunity to donate a share of their experimental earnings to Doctors Without Borders. Subjects then filled out a non-incentivized survey designed to measure socio-demographic variables and various behavioral traits that have previously been linked to behavior in coordination games. The questionnaire elicits patience (adapted from Andersen et al., 2008), reciprocity (adapted from Dohmen et al., 2009), risk preferences (adapted from Dohmen et al., 2011), trust (adapted from Fehr et al., 2003), cognitive ability (adapted from Frederick, 2005) as well as socio-demographic characteristics. We provide a more detailed description of the elicitation methodology in Section E in the Online Appendix. The post-experimental survey also included open-ended questions asking subjects about their motivation for how they behaved in the weak-link game. Experimental instructions and the post-experimental survey can be found in Section H of the Online Appendix.

The experiment was run at the laboratory facilities at the Centre for Experimental Economics (CEE), University of Copenhagen, in 2013 using the experimental software z-Tree (Fischbacher, 2007). The total duration of each session was approximately 45 min including reading the instructions, answering control questions, filling in the post-experimental questionnaire and picking up the payment. The main part of the experiment lasted about 15 minutes. The average earning of a subject was equal to 127.90 DKK (≈ $23.14) corresponding to an hourly wage rate of 170.53 DKK (≈ $30.86). For each session, the total and average donation and profit are shown in Table A3 in the Online Appendix.

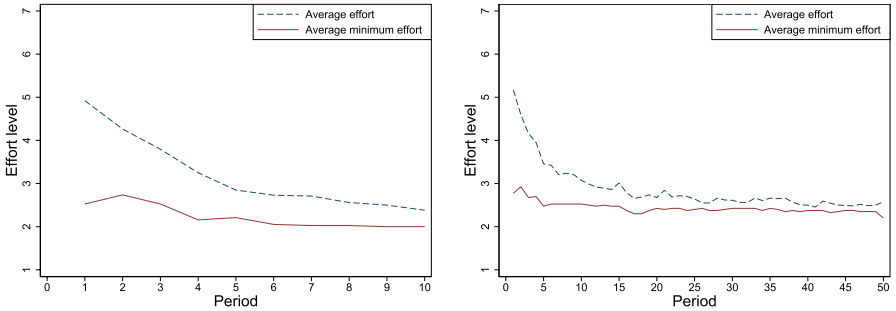
Before the start of the experiment, instructions about the experimental procedure were read aloud. If the number of participants exceeded a multiple of four, subjects were asked to leave the experiment for a show-up fee of 50 DKK.[7] To make sure that participants understood the rules of the coordination game and the corresponding payoff table, each subject was required to correctly answer three control questions. Players received payments for participation directly after the experiment.

Our experiment was conducted with 17–21 groups (68–84 subjects) per treatment, consequently resulting in a total of 78 groups (312 participants). Table A2 in the Online Appendix shows the number of subjects across the four treatments and eight sessions. Power calculations (reported in Section D of the Online Appendix) show that given our sample size and assuming a significance level of $\alpha = 0.05$, we could have detected the treatment effect in Berninghaus and Ehrhart (1998) of

[6] We encouraged subjects to make a choice on time by having a clock in red color on the screen that counts the remaining time before an effort choice should be made. After the allotted 15 s had passed, the subject saw a red blinking text in the right corner stating "Please reach a decision".

[7] In case there were not enough volunteers to leave the experiment, a randomly assigned card was handed out to the subjects indicating which participants are part of the coordination game and at what computer to work.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press



(a) Short horizon (10 repetitions) (b) Long horizon (50 repetitions)

Fig. 2 Average and group minimum effort across horizons

raising the number of periods from 10 to 30 ($\delta_M = 4.5$) with a power of more than 94%. Similarly, we could have detected the treatment effect of raising the number of periods from 10 to 90 ($\delta_L = 5.66$) with a power of more than 99.9%.

3 Experimental results

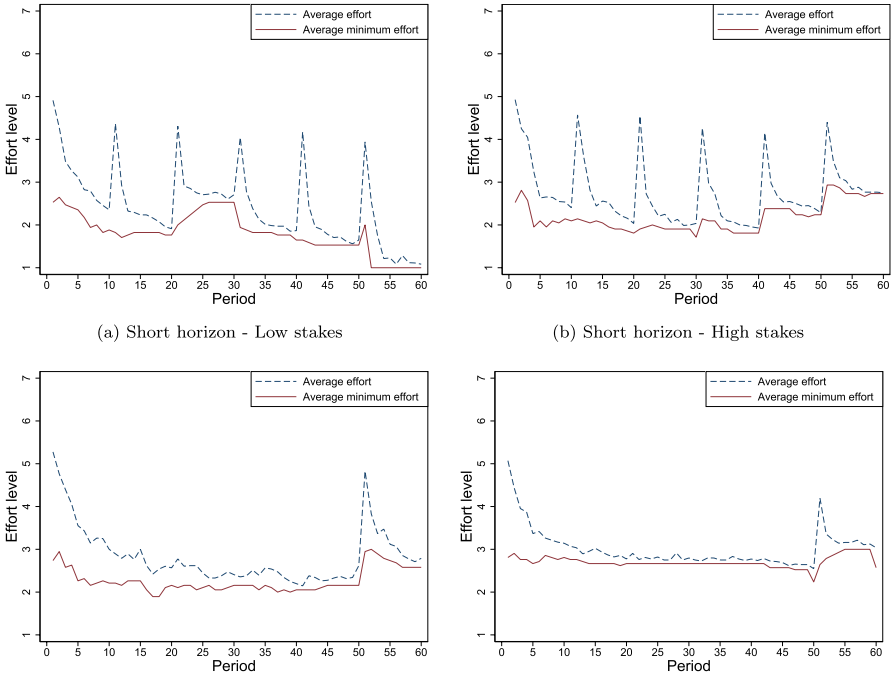
We first analyze the evolution of play graphically and the level of efficient coordination in the final period across our four treatments. We then proceed to analyze dynamic adjustments over time more carefully before turning to an econometric analysis that also investigates which socio-demographic factors and behavioral traits are conducive to coordination.

3.1 Main results

Figure 2 displays the average chosen effort levels and the average group minimum effort levels for the short (Fig. 2a) and long horizon (Fig. 2b) treatments. Figures A1 to A4 show the corresponding graphs for each group separately. Figure 2 reveals a typical pattern from earlier experimental studies: players initially choose intermediate effort levels, which gradually decline to low levels. There is no apparent difference between the short and long horizon treatments. 3 out of 38 groups in the short horizon treatments, and 4 out of 40 groups in the long horizon treatments, coordinate on the Pareto dominant equilibrium (see Table A4). This starkly contrasts the results reported in Berninghaus and Ehrhart (1998). In their experiment, 1 out of 6 groups in the short horizon treatment (10 repetitions), 3 of 6 in the medium horizon treatment (30 repetitions) and 5 of 6 groups in the long horizon treatment (90 repetitions) successfully coordinated on the Pareto dominant equilibrium.

Weber (2006, footnote 5) interprets Berninghaus and Ehrhart's (1998) findings as the result of a change of incentives, suggesting that the results are due to lower per-period stakes rather than a longer horizon. To analyze whether the per-period stakes play a role for coordination, Fig. 3 shows the average effort levels and average group minima for all four treatments separately. For full disclosure, Fig. 3 shows the results

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press



(a) Short horizon - Low stakes

(b) Short horizon - High stakes

(c) Long horizon - Low stakes

(d) Long horizon - High stakes

Fig. 3 Average and group minimum effort across treatments

also after the initial 10 or 50 rounds when subjects were rematched to new groups. Visual inspection of Fig. 3 reveals no apparent difference between the low and high stakes treatments. In the last period of the first round of weak-link games (i.e. period 10 in the short horizon and period 50 in the long horizon), 4 out of 36 groups in the low stakes treatment, and 3 out of 42 groups in the high stakes treatment coordinate on the Pareto dominant equilibrium.

Figure 3 also shows that all treatments display restart effects when groups are reshuffled every 10 periods in the short horizon treatments, as well as at the start of the last 10 periods (when groups were not reshuffled). The restart effects are always temporary and average effort levels decrease after a few periods. Table 3 shows some summary statistics for each of the four treatments and reports Kruskal–Wallis tests for differences between treatments. Table 3 indicates that there are no substantial differences in the best minimum effort level, the average minimum effort level and the average effort level across treatments either when comparing first-period choices, after 10 periods or after 50 periods.

The degree of coordination is affected both by initial effort choices and adjustments over time. Table 3 shows that players with low stakes tend to revise their effort choices more often. The difference is statistically significant when comparing high and low stakes in the long horizon treatments. To illustrate this difference, Fig. 4 shows the *activity level*, the percentage of players who change their effort level from one period to the next conditional on not already having reached the Pareto

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Table 3 Effort level and Kruskal–Wallis tests

	Treatment				Kruskal-Wallis test		
	Short horizon with low stakes	Short horizon with high stakes	Long horizon with low stakes	Long horizon with high stakes	*H*	*p*	*N*(Groups)
1 Period							
Best minimum effort level	2.53	2.52	2.74	2.81	0.59	0.9	78
Average minimum effort level	2.53	2.52	2.74	2.81	0.59	0.9	78
Average effort level	4.91	4.93	5.28	5.07	1.5	0.68	78
10 Periods							
Best minimum effort level	3.12	2.95	3.32	3.19	0.46	0.93	78
Average minimum effort level	2.22	2.23	2.43	2.79	0.69	0.88	78
Average effort level	3.21	3.19	3.81	3.69	3.09	0.38	78
% of changed effort level	52.0	42.3	54.0	44.6	5.92	0.12	78
50 Periods							
Best minimum effort level			3.32	3.19	0.17	0.68	40
Average minimum effort level			2.18	2.67	0.02	0.88	40
Average effort level			2.75	2.97	0.37	0.54	40
% of changed effort level			29.3	18.4	7.78	0.01	38

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

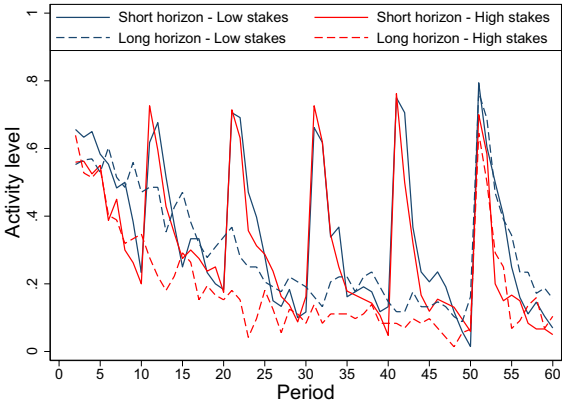


Fig. 4 Activity level across treatments

dominant equilibrium. (Again the figure shows all periods for full disclosure, but our statistical tests only use the initial 10 or 50 periods.) The fact that subjects appear to be more willing to experiment with low stakes could potentially affect coordination, but as noted above this difference does not result in different levels of coordination when stakes are lower.

To further investigate how players change effort levels over time, we follow Berninghaus and Ehrhart (1998) and separately analyze the learning directions of minimum and above-minimum players. Minimum players are here defined as subjects that choose $e_i^t = \underline{e}^t$, whereas above-minimum players choose an effort level above the minimum, or the maximum effort level if the group had already coordinated on the Pareto dominant equilibrium. Table 4 shows how minimum and above-minimum players adjust their strategies in the next period. A player can either reduce effort, keep effort unchanged or increase the effort. For comparison, Table 4 also displays the corresponding results from Berninghaus and Ehrhart (1998).

Table 4 reveals a striking difference between our results and Berninghaus and Ehrhart (1998). Whereas we find small differences across treatments, Berninghaus and Ehrhart (1998) report that minimum players are more likely to raise their effort level with a longer horizon, despite the fact that they used a shorter period length in treatments with a longer horizon and the last period's choice was the default. Berninghaus and Ehrhart (1998) also find that above-minimum players are less likely to reduce their effort choices with a longer horizon, which is perhaps less surprising given their default rule to implement the last period's choice. There is also a level difference between our study and Berninghaus and Ehrhart (1998). In their experiment, minimum players are much more likely to increase effort levels, and above-minimum players less likely to reduce effort levels. Consequently, subjects in our study have a stronger tendency to adjust to the group minimum rather than adjusting to above-minimum play. This tendency does not vary systematically with the number of repetitions, whereas it does so in Berninghaus and Ehrhart (1998). Because play in the initial period does not differ much compared to Berninghaus and Ehrhart (1998), the difference in the degree of coordination in the final period is likely to result from the differences in adjustments shown in Table 4.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Table 4 Learning directions

	Our treatment dimensions				Berninghaus and Ehrhart (1998)		
	Long horizon with high stakes	Long horizon with low stakes	Short horizon with high stakes	Short horizon with low stakes	Short Horizon	Medium Horizon	Long Horizon
Minimum players							
% of reduced effort levels	1	2.7	5.3	10.3	1.0	0.6	0.8
% of unchanged effort levels	90.6	78.0	76.3	65.5	51.3	31.3	37.1
% of raised effort levels	8.4	19.2	18.4	24.2	47.7	68.2	62.2
Non-minimum players							
% of reduced effort levels	70.6	51.4	62.0	60.6	42.5	30.9	20.7
% of unchanged effort levels	23.0	34.9	24.6	23.5	53.0	60.8	73.1
% of raised effort levels	6.4	13.7	13.5	15.9	4.6	8.3	6.2
N(Periods)	50	50	10	10	10	30	90
N(Groups)	21	19	21	17	6	6	6
N(Individuals)	84	76	84	68	48	48	48

Minimum players are defined as subjects that choose an effort level equal to the minimum group effort. Non-minimum players are defined as subjects that select an effort level above the minimum effort or the maximum effort level. The learning directions are based on the first 10 periods for the short and first 50 periods for the long horizon. In Berninghaus and Ehrhart (1998), the short horizon treatment consists of 10 periods, the medium horizon of 30 periods and the long horizon of 90 periods

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Table 5 Summary of behavioral preference measures

Preference	Item description	Weight
Risk aversion	Lottery choice sequence using the staircase method	0.5
	Self-assessment: willingness to take risk in general	0.5
Negative reciprocity	Self-assessment: willingness to take revenge	0.186
	Self-assessment: willingness to punish unfair behavior toward self	0.116
	Self-assessment: willingness to offend someone who offended	0.269
	Ultimatum game: minimum accepted amount	0.429
Altruism	Donation decision	0.5
	Ultimatum game: proposed offer	0.5
Trust	Self-assessment: general trust in people	0.318
	Self-assessment: ability to rely on other people	0.299
	Self-assessment: trust in strangers	0.383
Patience	Intertemporal choice sequence using the staircase method	1

The post-experimental questionnaire included open-ended questions asking subjects to motivate how they played in the experiment. We analyze their responses in Section G of the Online Appendix and we conclude that subjects' stated motivations align well with observed behavior. A substantial fraction stated that they deliberately played low numbers in the initial period and very few subjects reported that they tried to raise their effort level in the following periods.

3.2 Socio-demographics and behavioral traits

In this subsection, we examine what socio-demographic characteristics and behavioral traits are related to behavior in the weak-link game. The socio-demographic characteristics and behavioral traits are based on the general questions in the post-experimental survey (Section H.4), the Ultimatum Game (Section H.2) and the donation request (Section H.3). For each behavioral trait, the survey questions are combined into a single behavioral characteristic using the weighting procedure of indices suggested by Anderson (2008) and the variable classification applied in Falk et al. (2016) and Falk et al. (2018).[8] All non-binary items are transformed into standardized units at the individual level. The elicitation of the behavioral preferences and the construction of the preference indices are illustrated in Sections E and F of the Online Appendix. The behavioral measures and their corresponding weights are summarized in Table 5.

[8] Our survey module largely follows the experimentally validated approach measuring risk aversion, time discounting, trust, altruism and reciprocity in these papers. Based on an extensive experimental validation procedure, the weighted combination of these survey items is shown to perform best in predicting actual experimental behavior. The applied weights for the preference indices are quite similar to their proposed weights.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Table 6 Multiple regressions on individual first-period effort

	Final-period effort			
	(1)	(2)	(3)	(4)
Treatment				
High stakes	0.052	− 0.024	0.102	− 0.038
	[0.630]	[0.634]	[0.624]	[0.621]
Long horizon	0.265	0.120	0.243	0.053
	[0.645]	[0.654]	[0.663]	[0.668]
(High stakes) × (long horizon)	− 0.123	− 0.013	− 0.110	0.119
	[0.904]	[0.907]	[0.918]	[0.910]
Socio-demographics				
Danish		− 0.510*		− 0.648**
		[0.262]		[0.284]
Male		0.213		0.228
		[0.233]		[0.228]
Economics		− 0.116		− 0.135
		[0.238]		[0.248]
Experiment		− 0.027		− 0.037
		[0.250]		[0.258]
Game theory		− 0.059		− 0.001
		[0.259]		[0.256]
Age		0.042		0.008
		[0.151]		[0.152]
Parental education		− 0.058		− 0.098
		[0.136]		[0.138]
Cognitive ability		0.260*		0.271*
		[0.139]		[0.144]
Behavioral traits				
Risk aversion			0.011	0.039
			[0.123]	[0.127]
Negative reciprocity			0.006	− 0.068
			[0.109]	[0.119]
Altruism			0.047	0.048
			[0.101]	[0.101]
Trust			0.190*	0.258**
			[0.114]	[0.115]
Patience			− 0.157	− 0.176
			[0.123]	[0.128]
Treatment level	Individual	Individual	Individual	Individual
N(Individuals)	312	298	308	295
Period	10/50	10/50	10/50	10/50
Adjusted R^2	− 0.007	− 0.015	− 0.013	− 0.011

OLS regressions of the first-period effort level conditioning on the treatment variables (column 1), in addition to the socio-demographic factors (column 2), behavioral attributes (column 3) and all of the controls (column 4). "Economics", "Experiment" and "Game theory" refer to dummy variables indi-

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Table 6 (continued)

cating whether the subject reported having taken courses in economics, had previously participated in similar experiments or had taken classes in game theory. Significance levels of $p < 0.1$, $p < 0.05$ and $p < 0.01$ are indicated by */**/***. Heteroskedasticity-robust standard errors are reported within parentheses

We first regress the individual initial period effort level (e_i^1) on the socio-demographic factors, behavioral attributes and the treatment variables in the first period (Table 6). Unsurprisingly given the results presented above, the results of the regression analyses show that the two treatment variables, stake size and horizon, as well as their interaction are not pivotal for first-period effort choices. The fourth column of Table 6 indicates that the first-period effort choices are related to several socio-demographic and behavioral factors, but first period choices are most strongly related to cognitive ability. Higher effort choices are also negatively correlated with risk aversion, which is expected given that risk increases with the chosen effort level in the weak-link game. Trust, but not altruism, is positively correlated with initial effort choices. The association between being Danish and effort choices shown in column two of Table 6 is significant, but the regression coefficient becomes smaller and not significantly different from zero once behavioral traits are controlled for. This suggests that the high coordination success of Danes found in Engelmann and Normann (2010) might partly be mediated by differences in behavioral traits. For example, Danish subjects in our experiment exhibit statistically significantly higher levels of altruism, trust and patience, while showing a lower level of negative reciprocity.

We also analyze correlates of effort choices in subsequent periods in Section G of the Online Appendix. We find that risk averse subjects are less likely to play above the minimum, whereas more trusting players are more likely to do so. In contrast to first-period choices, there is no relationship between cognitive ability and above-minimum play. Moreover, whereas men tend to make lower first-period effort choices, they are more likely to play above minimum in subsequent periods. In a final set of analyses, we regress final-period group minimum on the group averages of the behavioral and socio-demographic variables. The strongest (and only statistically significant) predictor of final-period group minimum effort is the average cognitive ability of the group.[9]

4 Concluding remarks

Berninghaus and Ehrhart (1998) showed that coordination failures in the weak-link game can be overcome if players interact over many rounds. In their 90-period treatment, half of the groups managed to coordinate efficiently despite having played

[9] This is in line with previous experiments showing that cognitive ability is positively correlated with cooperative behavior (Baghestanian & Frey, 2016; Al-Ubaydli et al., 2016; Lohse, 2016; Jones, 2008; Proto et al., 2022), but contrasts with Al-Ubaydli et al. (2013) who do not find that cognitive ability is related to successful coordination.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

below efficient levels in the initial period. In our 50-period treatment, we only observe that play moves from sub-optimal levels to the efficient equilibrium in 2 out of 40 groups. Furthermore, we do not observe a difference when increasing the horizon while decreasing per-period stakes (as in Berninghaus & Ehrhart, 1998) or when holding per-period stakes constant. There are other design differences between the experiments that could potentially explain differences in results. For example, Berninghaus and Ehrhart (1998) vary the period length with the horizon and the previous period's decision is implemented by default when the time limit expires.

In a related study, Leng et al. (2018) let subjects play a weak-link game in (quasi-) continuous time, which can be thought of as a game with a very large number of repetitions. In their experiment, players can revise effort choices every 0.3 s during 60 s and then receive the accumulated flow payoff. The game is repeated 10 times with fixed groups. Leng et al. (2018) do not find higher degrees of coordination in continuous time compared to a control treatment in discrete time. To reconcile their findings with Berninghaus and Ehrhart (1998), Leng et al. (2018) note that substantial strategy revisions are primarily made at the beginning of each 60-s period. This restart effect means that it is primarily the beginning of each 60-s period that provides an opportunity to coordinate on a better equilibrium, so their continuous time game is perhaps better thought of as a 10-round game than a game with infinitely many rounds.

The lack of an effect of a longer horizon in our paper, and of the continuous time treatments in Leng et al. (2018), could potentially be explained by design differences compared to Berninghaus and Ehrhart (1998). However, together these results demonstrate that increasing the horizon or lowering per-period stakes is not a "quick fix" that always works to improve coordination in weak-link games.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40881-022-00125-z.

Acknowledgements We thank Erik Mohlin, Sten Nyberg, Jonathan de Quidt, David Strömberg and seminar participants at Stockholm University for helpful feedback. We are grateful for financial support from the Jan Wallander and Tom Hedelius Foundation.

Funding Open access funding provided by Stockholm School of Economics. Funding provided by Jan Wallanders och Tom Hedelius Stiftelse samt Tore Browaldhs Stiftelse (Grant number: P2017-0243:1)

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

References

Al-Ubaydli, O., Jones, G., & Weel, J. (2013). Patience, cognitive skill, and coordination in the repeated stag hunt. *Journal of Neuroscience, Psychology, and Economics, 6*(2), 71–96. https://doi.org/10.1037/npe0000005

Al-Ubaydli, O., Jones, G., & Weel, J. (2016). Average player traits as predictors of cooperation in a repeated prisoner's dilemma. *Journal of Behavioral and Experimental Economics, 64*, 50–60. https://doi.org/10.1016/j.socec.2015.10.005

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica, 76*(3), 583–618. https://doi.org/10.1111/j.1468-0262.2008.00848.x

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association, 103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Baghestanian, S., & Frey, S. (2016). GO figure: Analytic and strategic skills are separable. *Journal of Behavioral and Experimental Economics, 64*, 71–80. https://doi.org/10.1016/j.socec.2015.06.004

Berninghaus, S. K., & Ehrhart, K.-M. (1998). Time horizon and equilibrium selection in tacit coordination games: Experimental results. *Journal of Economic Behavior & Organization, 37*(2), 231–248. https://doi.org/10.1016/S0167-2681(98)00086-9

Bigoni, M., Casari, M., Skrzypacz, A., & Spagnolo, G. (2015). Time horizon and cooperation in continuous time. *Econometrica, 83*(2), 587–616. https://doi.org/10.3982/ECTA11380

Bryant, J. (1983). A simple rational expectations Keynes-type model. *Quarterly Journal of Economics, 98*(3), 525–528. https://doi.org/10.2307/1886025

Büyükboyaci, M. (2014). Risk attitudes and the stag-hunt game. *Economics Letters, 124*(3), 323–325. https://doi.org/10.1016/j.econlet.2014.06.019

Chaudhuri, A., Schotter, A., & Sopher, B. (2009). Talking ourselves to efficiency: Coordination in intergenerational minimum effort games with private, almost common and common knowledge of advice. *Economic Journal, 119*(534), 91–122. https://doi.org/10.1111/j.1468-0297.2008.02207.x

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1992). Communication in coordination games. *Quarterly Journal of Economics, 107*(2), 739–771. https://doi.org/10.2307/2118488

Cooper, R. W., DeJong, D. V., Forsythe, R., & Ross, T. W. (1990). Selection criteria in coordination games: Some experimental results. *American Economic Review, 80*(1), 218–233.

Crawford, V. P. (1991). An evolutionary interpretation of Van Huyck, Battalio, and Beil's experimental results on coordination. *Games and Economic Behavior, 3*, 25–59. https://doi.org/10.1016/0899-8256(91)90004-X

Devetag, G., & Ortmann, A. (2007). When and why? A critical survey on coordination failure in the laboratory. *Experimental Economics, 10*(3), 331–344. https://doi.org/10.1007/s10683-007-9178-9

Devetag, M. G. (2009). Coordination and trust as prerequisites of coopetition. In G. Battista Dagnino & E. Rocco (Eds.), *Coopetition strategy: Theory, experiments and cases* (pp. 274–289). London: Routledge.

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy, 91*(3), 401–419. https://doi.org/10.1086/261155

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2009). Homo reciprocans: Survey evidence on behavioural outcomes. *Economic Journal, 119*(536), 592–612. https://doi.org/10.1111/j.1468-0297.2008.02242.x

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association, 9*(3), 522–550. https://doi.org/10.1111/j.1542-4774.2011.01015.x

Engelmann, D., & Normann, H.-T. (2010). Maximum effort in the minimum-effort game. *Experimental Economics, 13*(3), 249–259. https://doi.org/10.1007/s10683-010-9239-3

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics, 133*(4), 1645–1692. https://doi.org/10.1093/qje/qjy013

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. IZA Discussion Paper.

Fehr, E., Fischbacher, U., Von Rosenbladt, B., Schupp, J., & Wagner, G. (2003). A nation-wide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative survey. CESifo Working Paper Series.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171–178. https://doi.org/10.1007/s10683-006-9159-4
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42. https://doi.org/10.1257/089533005775196732
Friedman, D., & Oprea, R. (2012). A continuous dilemma. *American Economic Review, 102*(1), 337–63. https://doi.org/10.1257/aer.102.1.337
Goeree, J. K., & Holt, C. A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review, 91*(5), 1402–1422. https://doi.org/10.1257/aer.91.5.1402
Goeree, J. K., & Holt, C. A. (2005). An experimental study of costly coordination. *Games and Economic Behavior, 51*(2), 349–364. https://doi.org/10.1016/j.geb.2004.08.006
Heinemann, F., Nagel, R., & Ockenfels, P. (2009). Measuring strategic uncertainty in coordination games. *Review of Economic Studies, 76*(1), 181–221. https://doi.org/10.1111/j.1467-937X.2008.00512.x
Jones, G. (2008). Are smarter groups more cooperative? Evidence from prisoner's dilemma experiments, 1959–2003. *Journal of Economic Behavior & Organization, 68*(3–4), 489–497. https://doi.org/10.1016/j.jebo.2008.06.010
Knez, M., & Camerer, C. (1994). Creating expectational assets in the laboratory: Coordination in 'weakest-link' games. *Strategic Management Journal, 15*(S1), 101–119. https://doi.org/10.1002/smj.4250150908
Kremer, M. (1993). The O-ring theory of economic development. *Quarterly Journal of Economics, 108*(3), 551–575. https://doi.org/10.2307/2118400
Leng, A., Friesen, L., Kalayci, K., & Man, P. (2018). A minimum effort coordination game experiment in continuous time. *Experimental Economics, 21*(3), 549–572. https://doi.org/10.1007/s10683-017-9550-3
Lohse, J. (2016). Smart or selfish-When smart guys finish nice. *Journal of Behavioral and Experimental Economics, 64*, 28–40. https://doi.org/10.1016/j.socec.2016.04.002
Neumann, T., & Vogt, B. (2009). Do players' beliefs or risk attitudes determine the equilibrium selections in 2x2 coordination games? FEMM Working Paper.
Proto, E., Rustichini, A., & Sofianos, A. (2022). Intelligence, errors, and cooperation in repeated interactions. *Review of Economic Studies, 89*(5), 2723–2767. https://doi.org/10.1093/restud/rdab095
Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review, 80*(1), 234–248.
Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *Quarterly Journal of Economics, 106*(3), 885–910. https://doi.org/10.2307/2937932
Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1993). Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication. *Games and Economic Behavior, 5*(3), 485–504. https://doi.org/10.1006/game.1993.1026
Van Huyck, J. B., Battalio, R. C., & Rankin, F. W. (2007). Evidence on learning in coordination games. *Experimental Economics, 10*(3), 205–220. https://doi.org/10.1007/s10683-007-9175-z
Weber, R. A. (2006). Managing growth to achieve efficient coordination in large groups. *American Economic Review, 96*(1), 114–126. https://doi.org/10.1257/000282806776157588

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

https://doi.org/10.1007/s40881-022-00125-z Published online by Cambridge University Press