# Learning together: behavioral, computational, and neural mechanisms underlying social learning in adolescence

Westhoff, B.

# 2

# Chapter 2

## Developmental asymmetries in learning to adjust to cooperative and uncooperative environments

*This chapter is published as:*
Westhoff, B., Molleman, L., Viding, E., van den Bos, W., & van Duijvenvoorde, A. C. K. (2020). Developmental asymmetries in learning to adjust to cooperative and uncooperative environments. *Scientific Reports*, *10*(1), 21761. https://doi.org/10.1038/s41598-020-78546-1

## Abstract

Learning to successfully navigate social environments is a critical developmental goal, predictive of long-term wellbeing. However, little is known about how people learn to adjust to different social environments, and how this behavior emerges across development. Here, we use a series of economic games to assess how children, adolescents, and young adults learn to adjust to social environments that differ in their level of cooperation (i.e., trust and coordination). Our results show an asymmetric developmental pattern: adjustment requiring uncooperative behavior remains constant across adolescence, but adjustment requiring cooperative behavior improves markedly across adolescence. Behavioral and computational analyses reveal that age-related differences in this social learning are shaped by age-related differences in the degree of inequality aversion and in the updating of beliefs about others. Our findings point to early adolescence as a phase of rapid change in cooperative behaviors, and highlight this as a key developmental window for interventions promoting well-adjusted social behavior.

## Introduction

Humans have evolved in a highly social environment in which they continuously make decisions about how to engage with others. Well-adjusted social behavior requires individuals to learn whom they can trust and cooperate with. We typically trust others whom we expect to reciprocate that trust in the future, and beliefs about others' trustworthiness are updated through everyday experiences. For example, if a friend violates our trust, this calls for an adjustment of our belief in their trustworthiness. This may not happen on the first violation, but if this friend continues their untrustworthy behavior, the friendship is unlikely to survive. Adjusting our beliefs based on outcomes of social interactions enables decision making that matches the situation and can be critical for successful navigation of the social world. In line with this notion, well-adjusted social behavior has been linked to positive developmental trajectories (e.g., in health, education, and social development), and is important for long-term mental health (Crone & Dahl, 2012; Dahl et al., 2018; Paus et al., 2008; Sawyer et al., 2018). With the rise of complex social worlds (both online and offline), learning about and adjusting to different social environments may be more important than ever. Yet little is known about the factors that underlie learning and adjusting in social environments, and how these skills manifest across adolescence.



Mounting evidence suggests that adolescence – the period between childhood and young adulthood – is a life phase in which learning and flexible behavior mature rapidly (see e.g., (Blakemore & Mills, 2014; Casey et al., 2008; Crone & Dahl, 2012; Sawyer et al., 2018). Moreover, adolescence is marked by a social reorientation: individuals start to form larger peer groups, peers gain in importance compared with parents, and social interactions become more complex (Chein et al., 2011; Larson & Richards, 1991; Larson et al., 1996; van Hoorn et al., 2016). Adolescence is, therefore, an important life phase for developing well-adjusted social behavior, with cooperative and uncooperative behaviors becoming more salient as adolescents deal with their social environments more independently. Developmental research into social decision making has shown that from an early age, children trust others and recognize that investing in others can lead to mutual benefits (Rosati et al., 2019). Cooperative behaviors, such as trust and prosocial behavior, are thought to continuously increase during adolescence (Blakemore & Mills, 2014; Eisenberg et al., 1991, 1995; van den Bos et al., 2010) (but see House et al., (2020)). A more nuanced view is that adolescents do not show more cooperative behaviors per se, but instead increasingly tailor their behavior to the social environment. For example, when undertaking prosocial actions they increasingly differentiate between friends and strangers (Fett, Shergill, et al., 2014; Güroğlu et al., 2014; van de Groep et al., 2020). Also, adolescents learn to adjust to interaction partners that differ in their level of trustworthiness (van den Bos, van Dijk, et al., 2012), something that children find difficult (Rosati et al., 2019).

Theoretically, decisions to cooperate can be based on at least three distinctive factors, each of which could differ between individuals and could contribute to developmental differences in adjusting behavior in social contexts: (1) social preferences, (2) prior expectations, and (3) updating of expectations. First, *social preferences* refer to individuals caring about relative outcomes, i.e. disliking having either a better or worse outcome than others (Fehr & Schmidt, 1999). Such social preferences show changes across development (see McAuliffe et al., (2017) for a review).The preference of avoiding getting *less* than others (i.e., disadvantageous inequality aversion) increases from age 4 to early adolescence (Fehr et al., 2008), which has been shown across many cultures (Blake et al., 2015), before it decreases again across adolescence (Meuwese et al., 2015). On the other hand, the preference of avoiding getting *more* than others (i.e., advantageous inequality aversion), appears from about 8 years old (Blake & McAuliffe, 2011; Fehr et al., 2008), and has been suggested to decrease across adolescence, particularly for boys (Meuwese et al., 2015). Nonetheless, also in adulthood people exhibit disadvantageous and advantageous inequality aversion (e.g., Dawes et al., 2007; Fehr & Schmidt, 1999). Although this evidence suggests that social preferences continue to develop across adolescence, little is known about how they impact learning in social environments. For instance, high levels of disadvantageous inequality aversion could prevent cooperative behavior due to a fear of getting less than others, even if there is a relatively strong expectation that others will cooperate.

Second, *prior expectations* (i.e., descriptive norms, the perceptions of what most people do (Chang & Sanfey, 2013; Delgado et al., 2005; Ruff & Fehr, 2014) inform decision making by generating predictions about the behavior of others (e.g., I will cooperate if this person is likely to reciprocate). Individual differences in initial expectations about others may lead to individual differences in choices (Fareri et al., 2015), and it is conceivable that different age groups have varying prior expectations (Ma, Westhoff, et al., 2020). However, prior expectations have hardly been studied in developmental populations, despite being important determinants of cooperative behaviors.

Third, adjusting behavior requires expectations to be *updated* in response to new information. Updating of expectations in social environments can be captured by reinforcement learning (RL) models (e.g., Cheong et al., 2017; Hackel et al., 2019; van den Bos et al., 2013), in which learning is driven by differences between expected and received rewards (i.e., prediction errors). Adolescence is characterized by substantial improvements in flexible learning and quick adaptation to novel non-social contexts (Decker et al., 2015; Hauser et al., 2015; van den Bos, Cohen, et al., 2012); whether this extends to the social domain, however, is still unclear (but see Rosenblau et al., (2018)).

Here we examine experimentally how children, adolescents, and adults adjust to social environments that differ in their level of cooperation, and aim to provide a mechanistic explanation by evaluating the role of social preferences, prior expectations, and expectation updating. To achieve this goal, we deployed a set of economic games, together with behavioral analyses and computational reinforcement learning modeling. Our cross-sectional sample spanned from late childhood into early adulthood (8 to 23 years old, *N*=244). Participants played age-appropriate versions of two well-studied incentivized economic games: A Trust Game (Figure 1B) and a Coordination Game (Figure 1D). These two games involve key types of cooperative behaviors: trust and coordination. Trust is key for mutually beneficial cooperation to be initiated and sustained (e.g., Berg et al., 1995; van den Bos et al., 2010; van den Bos, van Dijk, et al., 2012), and for achieving beneficial outcomes for all interaction partners involved. Yet, trust also creates a hazard of being betrayed. Similarly, coordinating one's behavior with others is often critical for collective welfare, even though outcomes may not always equally benefit all interaction partners (Luce & Raiffa, 1989; Osborne & Rubinstein, 1994).



The two games consisted of repeated one-shot interactions, in which both players had to choose between two options. In each trial, they encountered one new anonymous player from either a Cooperative environment or an Uncooperative environment in two games (the Trust Game and Coordination Game). The decisions of these players had been recorded in a previous session with age-matched unfamiliar others (see *Methods, pre-test*). Participants were explained that between environments, players could differ in their tendency to choose X (see Figure 1B and 1D). To maximize their earnings, participants had to learn over the course of the game which environment was Cooperative and which environment was Uncooperative, and adjust their choices accordingly. That is, in the Trust Game they had to learn in which environment the typical behavior was choosing X (labelled the 'Trustworthy environment') and in which environment the typical behavior was choosing Y (labelled the 'Untrustworthy environment'). Participants maximized their monetary outcomes by trusting (i.e., choose A) Trustworthy others and withhold trust (i.e., choose B) from Untrustworthy others. Similarly, in the Coordination Game, participants had to learn in which environment players tended to choose X (labelled the 'Friendly environment') and in which environment they tended to choose Y (labelled the 'Unfriendly environment'). Participants maximized their outcomes by coordinating with the response of the others, i.e., participants accepting a disadvantage (i.e., choose A) when interacting with the 'Unfriendly environment', or accepting an advantage (i.e., choose B) when interacting with the 'Friendly environment'. The social environments in these games were probabilistic, as cooperative behaviors were displayed by 73% of the players in the Cooperative environments, and by 27% of the players in the Uncooperative environments.

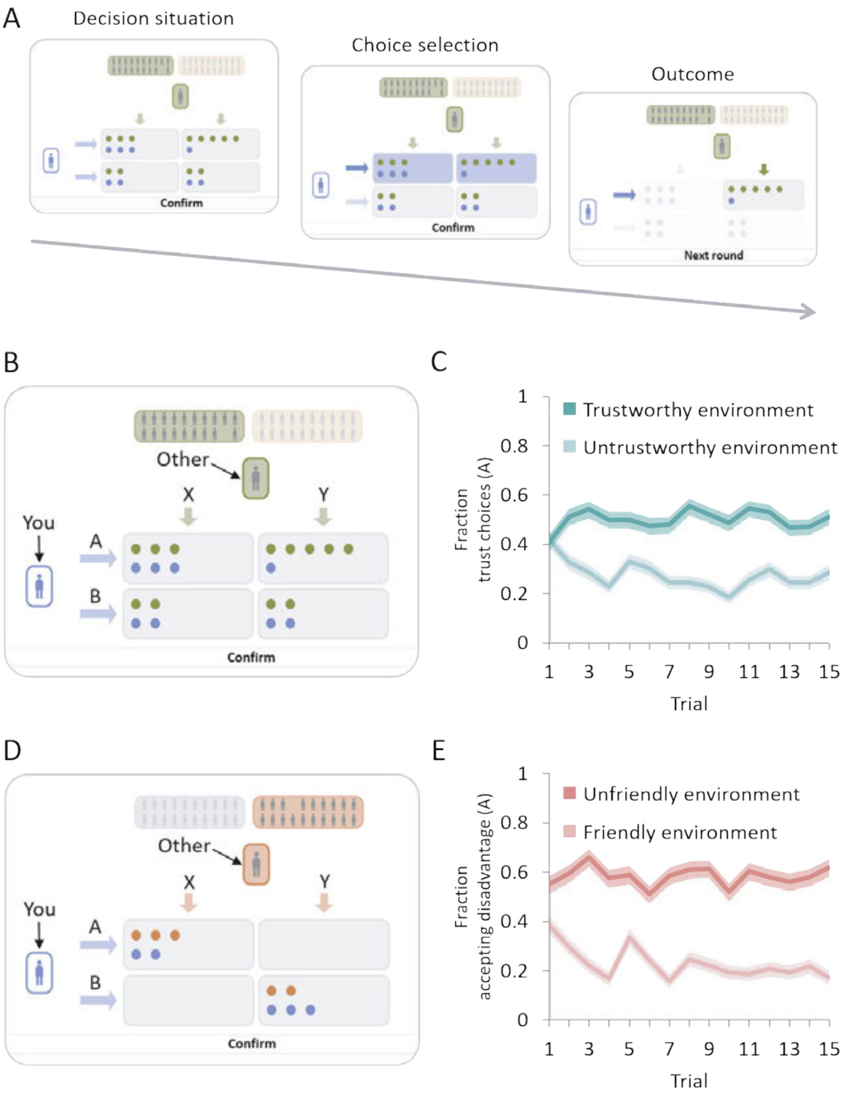



**Figure 1.** Task assessing learning to adjust to cooperative and uncooperative social environments. **(A)** Example trial. The participant (purple stick figure on the left) can choose between the top and bottom row of boxes (A or B). After choice selection, the participant is shown the pre-recorded choice (X or Y) of the other player (grey stick figure on the top). The background color of the other player indicates to which of the two environments they belong. The combined choices of the participant and the other player determine the monetary outcome for both players (number of dots in their corresponding color). **(B)** In the Trust Game, participants interact with players from a 'Trustworthy' environment (who tend to choose X) or an 'Untrustworthy' environment (who tend to choose Y). Participants' own monetary payoffs are maximized by choosing to trust (choose A) a player from a Trustworthy environment, and to withhold trust (choose B) from a player from the Untrustworthy environment. In this Trust Game

setup only disadvantageous inequality aversion may play a role in decision making. **(C)** Participant choices over trials per social environment, pooled across all participants. Over the course of the game, participants adjusted their choices by directing their trust towards players from the Trustworthy environment, and away from players from the Untrustworthy environment (N=244). **(D)** In the Coordination Game, participants' monetary payoffs are also maximized by matching the choices of their co-players. Again, social environments differ in their prevalence of (non)cooperation, reflected by players' tendencies to choose either X or Y. Coordinating on either of the outcomes (A,X) or (B,Y) will lead to positive outcomes. Respectively, participants' own monetary payoffs are maximized by choosing to accept a disadvantage (choose A) when confronted with a player from the Unfriendly environment, and to accept an advantage (choose B) when confronted with a player from the Friendly environment. In this game both advantageous and disadvantageous inequality aversion may play a role in social decision making. **(E)** Participant choices over trials per social environment, pooled across all participants. Overall, participants learned to coordinate with players from both environments. (N=202). Shaded areas in panels c and e represent standard errors of the mean (s.e.m.).



Participants also played an iterative Ultimatum Game (UG) and Dictator Game (DG), which allowed us to estimate participants' social preferences (i.e., advantageous and disadvantageous inequality aversion; see *Methods*). We separately assessed participants' prior expectations of the behavior of others before the start of the Trust Game and Coordination Game (see *Methods*). Furthermore, we used computational reinforcement-learning models (Daw et al., 2011) to model the updating of expectations between interactions. In these models, the learning rate quantifies how much an expectation violation modifies our subsequent expectations and consequently our decision making. We allowed learning rates to decay over the course of the games because we expected that most of the learning about the environments would happen in the first set of trials. After that, behavior would stabilize, provided the environments did not change their behavior (for more on learning rates and environmental stability see (Behrens et al., 2007; Li et al., 2011; Nassar et al., 2012)). We extended these reinforcement learning models to account for the measured prior expectations and social preferences (van den Bos et al., 2013), and compared the parameters of these models across age cohorts (see *Methods*).

We hypothesized that participants would be able to learn to adjust their behaviors to social environments differing in their level of (non)cooperation, but that across adolescence this ability would improve rapidly. We expected that these developmental differences could be explained by a combination of (1) social preferences (i.e., age-related changes in levels of advantageous and disadvantageous inequality aversion), (2) prior expectations (i.e., age-related changes in expectations about others' trustworthiness and tendencies to prioritize their own payoffs over those of others) and (3) updating of expectations (i.e., age-related changes in learning rates).

# Results

## Learning to adjust to cooperative and uncooperative social environments across age

First, we examined decisions over the course of the games to assess whether children, adolescents, and young adults adjust their behavior to different social environments with different levels of cooperation. For this, we used the Trust Game in which participants maximized their monetary outcomes by trusting Trustworthy others and withhold trust from Untrustworthy others (Figure 1B), and the Coordination Game in which participants maximized their outcomes by coordinating with the response of the others, i.e., participants accepting an advantage when interacting with the 'Friendly environment', or participants accepting a disadvantage when interacting with the 'Unfriendly environment'. We performed a binomial generalized linear mixed model (GLMM) per game on participants' binary choices, including social preferences and prior expectations of others' behavior (see *Methods*).

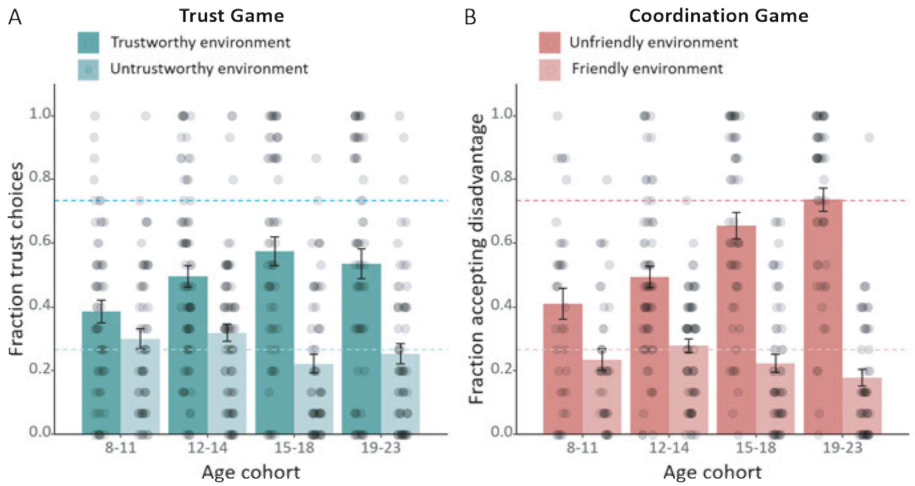



**Figure 2.** Developmental asymmetries in adjusting to cooperative and uncooperative social environments. Decisions per age cohort are shown per social environment in the **(A)** Trust Game (*N* = 244) and **(B)** Coordination Game (*N* = 202). Across age cohorts, there is a pronounced increase in cooperative choices (i.e., trusting players from a Trustworthy environment in the Trust Game, and accepting a disadvantage from players from an Unfriendly environment in the Coordination Game), whereas uncooperative choices (i.e., withholding trust from players from an Untrustworthy environment in the Trust Game, and accepting an advantage from players from a Friendly environment in the Coordination Game) were relatively stable across age. Dashed lines in a and b indicate the reinforcement rates (i.e., fraction of 0.73 and 0.23) for each social environment. Note that we grouped participants for illustration purposes only and age was treated as a continuous variable in all analyses. Error bars represent s.e.m.

For the Trust Game, results indicated an accelerated change in adolescence in which people differentiated more between the Trustworthy and Untrustworthy environment (environment x age linear, $B = -0.307$, $p < .001$; environment x age quadratic, $B = 0.205$, $p = .015$; N=244; see Table S1 for full statistical analysis; Figure 2A). Post-hoc tests per social environment showed that trusting the Trustworthy others increased rapidly in early to mid-adolescence (age linear, $B = -0.384$, $p = .006$; age quadratic, $B = 0.316$, $p = .020$). In contrast, adjusting to Untrustworthy others improved slightly, and monotonically across adolescence (age linear, $B = 0.233$, $p = .031$).

For the Coordination Game, results again indicated that with age, people differentiated more between the Friendly and Unfriendly environment (environment x age linear, $B = -0.458$, $p < .001$; $N = 202$; see Table S3 for full statistical analysis; Figure 2B). Post-hoc tests per social environment showed that optimally coordinating to the Unfriendly environment (i.e., participants accepting their disadvantage) increased across adolescence (age linear, $B = -0.446$, $p < .001$). However, coordinating to the Friendly environment (i.e., participants accepting their advantage) did not change with age; participants from all age cohorts adjusted quickly to this environment. Together, these results show that people coordinated to both environments but younger participants were less likely to accept a disadvantage than older participants.

## Social preferences and prior expectations

Social preferences (advantageous and disadvantageous inequality aversion) and prior expectations of others' behavior are features that may account for age-related changes in learning to adjust to different social environments. Before further testing their relation to behavior in the Trust Game and Coordination Game, we first examined the age-related changes in these parameters. Robust linear regression analyses (5000 bootstraps) indicated that only disadvantageous inequality aversion changed across age (Figure 3A-3D). Specifically, older participants were, compared to younger participants, less averse to being behind (age linear, $B = -0.098$, $\beta = -0.308$, $p < .001$, 95% CI = [-0.139, -0.057], $N = 244$). We did not observe significant age-related change for advantageous inequality aversion (age linear, $B = -0.099$, $\beta = -0.118$, $p = .133$, 95% CI = [-0.229, 0.031], $N = 202$), nor for prior expectations of others' trustworthiness (age linear, $B = -0.048$, $\beta = -0.085$, $p = .209$, 95% CI = [-0.124, -0.027], $N = 245$) or for prior expectations of others' tendency to prefer to have more than the other (age linear $B = -0.031$, $\beta = -0.076$, $p = .209$, 95% CI = [-0.087, -0.024], $N = 245$).

In a binomial GLMM analysis, advantageous and disadvantageous inequality aversion were related to choices in the games (see Tables S1 and S3 for full statistical analysis). Greater disadvantageous inequality aversion was associated with overall fewer trusting choices ($B = 0.215$, $p = .012$) and with fewer choices in which participants accepted a disadvantage ($B = 0.134$, $p = .048$). In addition, greater advantageous inequality aversion was associated with greater acceptance of a disadvantage (*B* = 0.172, *p* = .009). In contrast, prior expectations were not related to choices in both games.

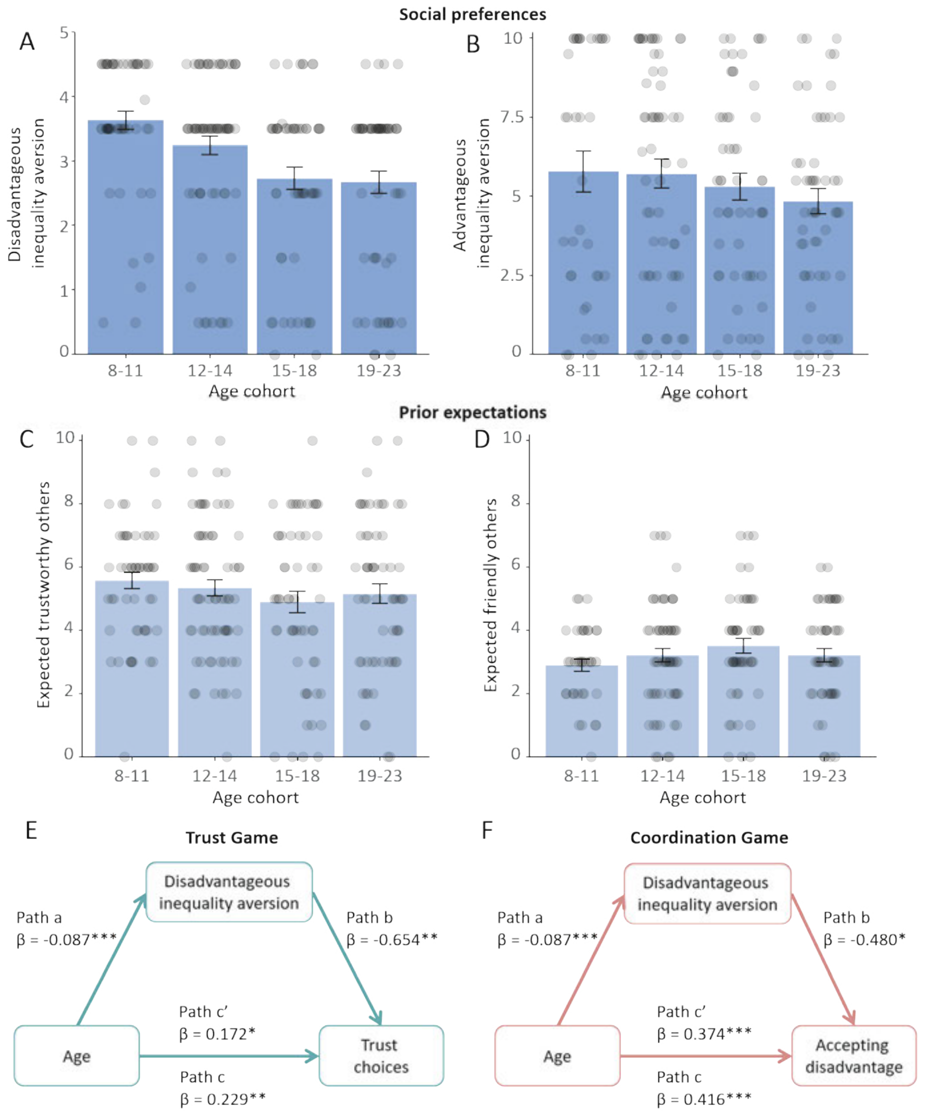



**Figure 3.** Social preferences and prior expectations across age cohorts. Social preferences and prior expectations are features that may account for choices in the games. **(A)** Indifference points of the Ultimatum Game as a measure of disadvantageous inequality aversion and **(B)** Indifference points of the Dictator Game as a measure of advantageous inequality aversion were used as social preference measures. **(C)** Prior expectations for the Trust Game, with higher values indicating greater expectations

that others are Trustworthy. **(D)** Prior expectations for the Coordination Game, with higher values indicating greater expectations that the others are Friendly (i.e., not tending to prioritize their own payoffs over those of others). In panels a-d, error bars show s.e.m. **(E)** Mediation model for the effect of age on trust behavior towards players from the Trustworthy environment (controlled for trust choices towards players of the Untrustworthy environment), via disadvantageous inequality aversion. **(F)** Mediation model for the effect of age on coordination behavior (i.e., accepting a disadvantage) when confronted with players from the Unfriendly environment (controlled for accepting disadvantages when playing with the Friendly environment), via disadvantageous inequality aversion. Note in mediation models: c = total effect, c' = direct effect; values are standardized regression coefficients of direct effects, and asterisks indicate significance levels (*$p < .05$, ** $p < .01$, *** $p < .001$).



To better understand what drives the age-related change in learning to adjust to the social environments differing in their level of cooperation, we ran a mediation analysis per game. Specifically, we examined whether the age-related changes the observed increase in cooperative behavior (trusting, accepting a disadvantage) across development, was explained by the age-related change in disadvantageous inequality aversion (Figure 3E; Figure 3F). We found that, when controlled for choice participants' choice behavior in the Untrustworthy environment, the improvement across age in adjusting to the Trustworthy environment ($\beta = 0.172$, $p = .026$) was partly explained by the age-related decrease in disadvantageous inequality aversion (indirect effect = 0.057, $SE = 0.026$, 95%CI = [0.011, 0.113]). That is, older participants showed lower levels of disadvantageous inequality aversion ($\beta = -0.087$, $p < .001$), which in turn resulted in more trust choices ($\beta = -0.654$, $p = .007$; Figure 3E). This mediation analysis for the Coordination Game similarly showed that, when controlled for choice behavior in the Friendly environment, the age-related improvement in coordinating with the Unfriendly environment ($\beta = 0.374$, $p < .001$) was partly explained by disadvantageous inequality aversion (indirect effect = 0.042, $SE = 0.020$, 95% CI = [0.005, 0.083]). That is, older participants showed lower levels of disadvantageous inequality aversion ($\beta = -0.087$, $p < .001$), which in turn resulted in more acceptance of a disadvantage when coordinating with the Unfriendly environment ($\beta = -0.480$, $p = .022$; Figure 3F). Note that this partial mediation in the Coordination Game did not hold when advantageous inequality aversion was included as an additional mediator.

## Computational modeling of updating expectations

To understand how children, adolescents, and young adults update their expectations in different social environments, we developed computational models that extend basic reinforcement learning models(Rescorla & Wagner, 1972). In our models, participants use the outcome of interactions to update their expectations of their interaction partners' choices in each social environment (Figure 4A-C). The extent to which these expectations are updated is reflected in a learning rate ($\lambda$). Besides quantifying the updating of expectations, this computational approach allows us to confirm the role of social preferences as observed in our behavioral analyses. We extended the basic reinforcement model by i) incorporating mean cohort-level social preferences to calculate a subjective value of interaction monetary outcomes (Figure 4A, 4C) that drives decision making, and ii) by allowing learning rates (expectation updating) to exponentially decay over trials of the game. Thus, we fitted four variants of this model (with and without social preferences; with and without decaying learning rates) to our experimental data for each age cohort and each game, to allow estimating different parameters (learning rates, expectation updating) across cohorts per game (see *Methods*).

For both the Trust Game and the Coordination Game, a comparison of model fits provided strong support for models extended with social preferences (Figure 4D), confirming the results from our behavioral analyses that social preferences impact decision making. The best models also included decaying learning rates (see Figure 4D and Table S7). For the Trust Game, we observe that for the 8-11 year-olds, estimated learning rates are constant over the course of the game, suggesting that in late phases, individuals in this youngest age cohort still updated their expectations of the behavior in the different social environments. In the older age cohorts, learning rates start high (around $\lambda$=1; asymptote not shown in Figure 4E) and decay over trials, indicating that expectations take form relatively early in the game, and remain relatively stable later on. For the Coordination Game (Figure 4F), we observe a similar pattern: older participants tended to show the strongest decay in learning rates over trials, whereas participants from the younger cohorts tended to update their expectations throughout the game.

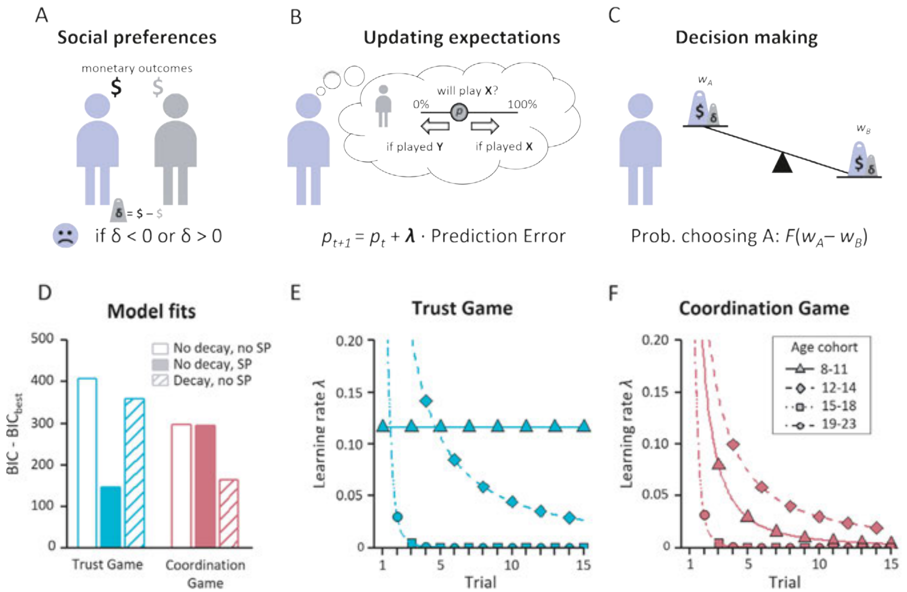



**Figure 4.** Model of reinforcement learning with social preferences. **(A)** Individuals acquire monetary payoffs ($) from interactions in the games (*cf.* Figure 1). How they subjectively value each outcome can be influenced by their social preferences (Fehr & Schmidt, 1999), whose impact is proportional to the differences in payoff between interaction partners (δ). **(B)** Over the course of the games, individuals update their expectations ($p$) about the behavior of the other players. The extent of updating is proportional to the prediction error (i.e., the difference between the expected and actual outcome) and learning rate $\lambda$ (Sutton & Barto, 1998). **(C)** The probability that an individual chooses A (not B) is an increasing function ($F$) of the difference in weight of the two choice options ($w_A$ and $w_B$; Methods). We use maximum likelihood methods to estimate learning rates $\lambda$ for each age cohort and each game separately. In the baseline model, the weights reflect individuals' monetary payoff from their interactions ($) and $\lambda$ is constant over trials. Extended models also incorporate social preferences and learning rates that decay exponentially over game trials (see *Methods*). **(D)** For both games, bars show BIC differences of three model variants with the best model, which includes both social preferences (SP) and decaying learning rates. For each model, BIC values were calculated by summing BIC values of all age cohorts. Solid bars reflect models including social preferences; void and hatched bars respectively show models excluding and including decaying learning rates. **(E)** Estimated learning rates from the best models as a function of the trial number, for each of the four age cohorts separately for the Trust Game, and **(F)** for the Coordination Game. For the older cohorts (15-18 and 19-23), the estimated learning parameters were virtually identical, causing the plotted lines to coincide.

## Discussion

Here, we examined children's, adolescents' and adults' ability to learn to adjust to social environments that differ in their level of cooperation. We examined the role of social preferences (inequality aversion), prior expectations about others' behavior, and the updating of expectations as potential mechanisms underlying this behavior. To this end, participants played a series of economic games with groups of age-matched unfamiliar others, which captured two important cooperative behaviors: trust and coordination behavior. Our results show a striking developmental asymmetry in the learning to adjust to (un)cooperative environments: people adjust well to environments that require uncooperative behavior (i.e., withholding trust, accepting an advantage) from a young age, yet only during adolescence they learn to adjust to environments that require cooperative behaviors (i.e., trusting, accepting a disadvantage). Our results provide several insights into the mechanisms that explain these age-related differences.

First, age-related differences in learning to adjust to cooperative behaviors can be partly explained by differences in social preferences. Specifically, older participants showed lower levels of disadvantageous inequality aversion which explained their higher levels of cooperative behaviors in a Trust Game and Coordination Game. That is, younger participants are less willing to cooperate (trust, accept a disadvantage) given that they are more averse to potential non-cooperation of the other player. Moreover, our computational models confirmed that participants' decisions were best captured by a reinforcement learning model extended with social preferences in all age bins. Note that in our current RL modeling approach social preferences influence choice behavior through a subjective transformation on the participants' expected payoffs in these games. This way, the RL models show e.g., how disadvantageous inequality aversion (dislike of being behind) can reduce the ability to adjust to environments that require cooperative behaviors. Future work should explore whether other factors underlie the age-related changes in social preferences and their mediating effects on cooperative (adjusting) behaviors. For example, the willingness to punish disadvantageous outcomes, or trying to force the other to coordinate in your favor may be alternative motivations underlying these effects. However, those questions are better answered by using experimental designs in which participants play multiple rounds with the same person, rather than a series of one-shot games. Taken together, our results underline that for understanding age-related changes in social decision making it is critical to understand the development in social preferences, which differ across developmental windows and largely drive social decision making.

A potential mechanism that may relate to the influence of inequality aversion on decision making, is behavioral control (McAuliffe et al., 2017). Behavioral control refers to the ability to control thoughts and actions in order to regulate behavior towards (long-term) goals (Miller & Cohen, 2001; Steinbeis, 2018b). Developmental studies have shown that behavioral control undergoes protracted development due to a prolonged maturation of underlying neural circuitry in regulatory brain regions including the prefrontal cortex (Achterberg et al., 2016; Steinbeis, 2018b; van den Bos et al., 2015). In turn this would result in developmental changes in responses to inequality into childhood and presumably into adolescence (Meuwese et al., 2015; Sul et al., 2017). An experiment in children also confirmed a direct role of behavioral control in behavior that benefits others: taxing children with a response inhibition task resulted in less prosocial behavior and more costly punishment to violations of fairness (Steinbeis, 2018a). An alternative explanation is that inequality may evoke stronger emotional responses, such as increased levels of anger (Güroğlu et al., 2011; van den Bos, van Dijk, et al., 2012). This would yield a different view on social preferences in which responses to inequality can be based on emotion regulation ability. Future studies are necessary to further disentangle whether such self-regulatory behaviors drive the development of social preferences and their influence on cooperative behaviors. Consequently, an interesting field for future studies is whether strengthening self-regulatory processes is a promising pathway for stimulating co-operative behavior in young people.



Besides social preferences, we also examined how people's *prior expectations* of others' trustworthiness and inclination to take more than others influenced learning in different social environments. Our results indicated that reported prior expectations of others' behavior were stable across age cohorts. This is surprising given the consistently reported increase in cooperativeness across age (e.g., Blakemore & Mills, 2014; Eisenberg et al., 1991, 1995; van den Bos et al., 2010), which was also observed in the current experiment. This suggests that there is a developmental mismatch between prior expectations and the actual levels of co-operation. Moreover, contrary to our hypotheses, we did not find effects of prior expectations on learning to adjust to different social environments. Perhaps people do not have strong prior expectations about others' behavior in the anonymous games used in the current study, and any expectations they might have are overridden quickly by outcomes of interactions. Presumably, effects of prior expectations in the current setup would be more prominent in a more heterogeneous sample with greater diversity in - for example - life-history backgrounds. For instance, prior expectations (as well as the updating of these expectations), may be different for people who have grown up in an environment where rewards and punishments are unpredictable, this may be particularly the case for children who have experienced harsh and inconsistent discipline, maltreatment and neglect (Hanson et al., 2017; Pitula et al., 2017). These expectations of others' behavior may match their environmental experiences and as such, they may engage in social situations differently. Thus, when assessing the generalizability of our results it would be important to include a more heterogeneous sample with greater diversity in life-history backgrounds. Including different populations could also help answering the question to what extent prior expectations about behavior in games reflect prior expectations about cooperative behavior in the real world (e.g., Benz & Meier, 2008).

Here we used computational modeling to quantify how quickly children and adolescents updated their expectations based on choice outcomes in previous interactions. Interestingly, when placed in a new social environment, people were initially highly sensitive to behaviors of other players, and quickly adapted their behavior to the outcomes they experienced. For older ages, behavior stabilized after a few interactions as signaled by a decrease in learning rate. Children and young adolescents, however, continued to react to the choices of others across the games. That is, they often switched strategies after a surprising response from one of the environments. This finding indicates that during adolescence, people more effectively integrate outcomes over time, and consequently form stable expectations of others based on their behavior, which are not quickly overridden by a single experience. Building lasting relations may crucially depend on this integrated information of others' behavior. Although the continuous expectation updating of children and adolescents hampers their learning in stable environments, this actually may provide an advantage in fast-changing or unpredictable environments (Nussenbaum & Hartley, 2019). That is, in such environments, immediately responding to changing feedback is more beneficial than sticking to prior expectations (Decker et al., 2015). Whether fast-updating better fits children's and adolescents' experienced social environments is an interesting question for future studies.

In the current study, participants were confronted with choices from actual peers and real-life consequences of their actions for all interaction partners. This two-directional approach, rather than often-used one-way decision making, is acknowledged as an important aspect of paradigms in social sciences (Camerer & Mobbs, 2017). However, the controlled social environments in our study are less complex than real-life social interactions, in which factors such as social status, culture, or reputation may complicate social decision making. Future studies, e.g., field studies or studies using virtual reality, could aim to further approach the complexity of real-life social interactions, while retaining experimental control. In addition, we included a specific experimental set-up of social learning in which participants were given prior information on the different social environments. Future studies will need to assess whether our developmental findings hold in settings where participants need to figure out base rates of cooperativeness and exploitation on their own. Another limitation of the current study is that whereas social preferences were revealed preferences, prior expectations were stated expectations about others. People find it hard to estimate probabilities, and future studies need to assess the validity of these preferences with individual difference measures. Moreover, although IQ did not differ between groups and did not influence any of our findings, our adult participants were mainly recruited through university advertisements. Future studies should aim for a representative sampling strategy in each age cohort. A final limitation of the current study is its cross-sectional design, as longitudinal studies are necessary to identify developmental patterns. Therefore, developmental interpretations of behavioral results and the underlying mechanisms remain speculative.

In sum, we combined computational learning models and experimental social manipulations to demonstrate age-related changes in adjusting cooperative behaviors. Well-developed social skills are essential for succeeding in society and long-term positive outcomes. The ability to adapt to different social environments and discern who we should trust and cooperate with, may benefit short-term outcomes, but may also foster social relationships and restrain behavioral and mental health problems in the long-term (Crone & Dahl, 2012; Dahl et al., 2018; Paus et al., 2008; Sawyer et al., 2018). Knowledge of how such social adjustment behavior manifest in different developmental stages inform what ages are the important developmental phase for monitoring social development, and what ages are potentially more receptive to interventions (Dahl et al., 2018; Yeager et al., 2018).



Our study has shown that adjusting cooperative behaviors is developing rapidly in early adolescence. Improvements in adjustment to different social environments are driven by developing social preferences (waning aversion to disadvantageous inequality aversion) and increasingly effective updating of own behavior in response to others' behavior. Early adolescence would, therefore, be a key target window for interventions targeted at stimulating cooperative and well-adjusted social behavior. Moreover, these findings provide important starting points for interventions for youth with maladaptive social tendencies, such as youth with conduct disorder problems(Frick & Viding, 2009; Viding & McCrory, 2019).

## Methods

### Participants

A total of 269 participants (58.4% female) between ages 8 and 23 years took part in this study. Participants were recruited from a primary school (*n* = 60), two secondary schools (*n* = 128), and through local advertisements at a university campus (*n* = 81) in the western and middle part of The Netherlands. The majority of the participants (92.3%) were born in the Netherlands, and a minority was born elsewhere (Morocco 1.4%; all other countries <1%), or information was missing (1.4%). Twenty participants from secondary schools (ages 14-16) were excluded due to technical problems with saving the learning data. Four participants were excluded because they did not finish the cognitive behavioral measures, and therefore IQ could not be estimated. The final sample consisted of 245 individuals aged between 8 and 23 years.

Adult participants provided written informed consent. For minors, written informed consent was obtained from parents. To make the tasks incentive-compatible, participants were informed that with each behavioral task they could win points that represented lottery tickets. In each class, and in a similar-size group of adults, one lottery ticket was randomly drawn and the winner received a digital 10 Euro gift voucher. In addition, all minors received a small gift; adults received 10 Euros flat rate or course credit. All procedures were approved by the Psychology Research Ethics Committee of Leiden University (minors: CEP17-0301/120;

young adults: CEP17-1009/334) and performed in accordance with the relevant guidelines and regulations.

For analyses using age cohorts (see *Computational modeling* in the section below), we divided the sample into four roughly equally-sized age cohorts: 8-11 year-olds (n=54, 46.3% female, mean age 10.6, SD 0.9), 12-14 year-olds (n=73, 52.1% female, mean age 13.4, SD 0.7), 15-18 year-olds (n=57, 59.6% female, mean age 17.0, SD 1.3), and 19-23 year-olds (n=61, 80.3% female, mean age 21.1, SD 1.4). A $\chi^2$-test indicated sex differences between age cohorts ($\chi^2_{(3)}$ = 16.6, $p$ = .001), with more females in the oldest age cohort. IQ was estimated using a speeded version of the Raven Standard Progressive Matrices (Hamel & Schmittmann, 2006). The estimated IQ scores were largely within the normal range varying between 79 and 136 (mean IQ = 106, SD = 10.3), and did not differ significantly between age cohorts ($F$ (3,237) = 2.18, $p$ = .090) and sexes ($F$ (1,237) = 0.28, $p$ = .770). Additional analyses showed that sex differences and IQ did not confound performance on the social games, and did not influence any of our observed age-related changes therein (see Tables S2 and S4).

## Pre-test

A key component of the economic games used in the current study is that choices have consequences not only for oneself, but also for the other player. To ensure this, we performed a pre-test at a separate high school and a separate adult sample (both in The Netherlands) functioning primarily as a match for determining the participants' outcomes and thereby creating a true social consequence of behavior.

In total, 82 adolescents and 44 adults were asked to make one choice (X or Y) for each social game (Trust Game and Coordination Game, see Figure 1). We randomly linked each participant in the full-experiment with one pre-test participant. This match and the combined outcomes of their choices determined the outcome for the participants (number of points), as well as for the pre-test participant. The pre-test participants had a similar lottery ticket procedure as the participants from the full experiment, i.e., points were lottery tickets with which they had a chance of winning a 10 Euro gift voucher. All pre-test participants received a similar instruction as the participants of the main study. That is, it was stressed that their choices would have consequences for themselves and another participant, since their outcomes would result from their combined choices.

## Economic games: Trust Game and Coordination Game

Participants completed two incentivized economic games: A Trust Game and a Coordination Game (Figure 1). Each game was composed of 30 trials in total: each trial was a one-shot game with a new anonymous player (whose decision had been recorded in the pre-test; see above). Every trial the participants chose between 2 options (A or B) to distribute points between themselves and the other. After their decision they could see the choice of the player

(X or Y) and the outcomes for themselves and the player. Outcomes for self and the player resulted from their combined choices, as shown with payoff matrix $\begin{bmatrix} \boldsymbol{a}, a' & \boldsymbol{b}, b' \\ \boldsymbol{c}, c' & \boldsymbol{d}, d' \end{bmatrix}$ where in each of the cells entries with and without apostrophes indicate payoffs for, respectively, the other and self (in bold).

In each of the games, the two social environments consisted of 20 players each (but note that participants interacted with only 15 players per environment). Environments are formed based on pre-test responses, which were matched to create a 'Cooperative' (73%, i.e., 11 out of 15) and an 'Uncooperative' social environment (Figure 1). Over the course of the game trials, participants could learn the tendency of choosing X for each environment of other players, and adjust their responses accordingly. Participants were incentivized by associating their performance to the chance of winning a gift voucher (see Supplementary materials for the instruction protocol).



The Trust Game (Figure 1B) was characterized by payoff matrix $\begin{bmatrix} \boldsymbol{3}, 3 & \boldsymbol{1}, 5 \\ \boldsymbol{2}, 2 & \boldsymbol{2}, 2 \end{bmatrix}$. Participants could maximise their earnings by choosing A ('trust'; top row) when matched with a member of the Trustworthy environment, and choosing B ('not-trust'; bottom row) when matched with a member of the Untrustworthy environment. The Coordination Game (Figure 1D) was characterized by payoff matrix $\begin{bmatrix} \boldsymbol{2}, 3 & \boldsymbol{0}, 0 \\ \boldsymbol{0}, 0 & \boldsymbol{3}, 2 \end{bmatrix}$. Participants could maximize their earnings by coordinating to their partners' choices. That is, the participant needed to accept a disadvantage (choose A; top row) when matched with a member of the 'Unfriendly' environment, but when matched with a member of the 'Friendly' environment the participant needed to accept an advantage (choose B; bottom row).

The order of these two games was counterbalanced across participants. Within each game, participants played 30 trials, 15 trials with each environment of players (e.g., Trustworthy and Untrustworthy environment). The inconsistent choices within an environment (e.g., Y when playing with someone of the environment that prefers X) were distributed across trials, yet fixed on trials 4, 8, 12, and 14. Within a game, the order of interactions with the two different environments was presented randomly, yet fixed across participants.

Although our main research questions center on the factors specific to learning to adjust behavior in different social environments (e.g., the role of prior expectations about others, and getting more or less than others), we also included a non-social learning task to examine the level of behavioral adjustment in a simple learning context (Figure S1 and Tables S5-S6). In this non-social learning task, participants played with computers as interaction partners, and only the participant – not the computers – could receive payoffs. A formal comparison between age-related changes in learning to adjust to non-social versus social environments is included in the Supplementary materials. A computational modeling approach on the non-social learning task is discussed in Figure S5.

## Social preferences

We measured advantageous inequality aversion and disadvantageous inequality aversion in two separate tasks: respectively, a modified Dictator Game (DG) and Ultimatum Game (UG). These measures were derived from an adapted (i.e., child-friendly and short) version of a DG and UG (based on (Beranek et al., 2015; Blanco et al., 2011)). Participants always performed the DG and UG right before the economic games.

In the Dictator Game participants were given six binary choices to divide 10 points between themselves and another anonymous participant in the study; one option was always an unequal distribution (10/0; 10 points for self, 0 points for the recipient) and the other option an equal distribution of points for themselves and the recipient (i.e., starting with (5, 5) and decreasing to (0, 0) with each subsequent trial [ (4, 4), (3, 3), (2, 2), (1, 1), (0, 0)] or increasing to (10,10) with each subsequent trial [ (6, 6), (7, 7), (8, 8), (9, 9), (10, 10)], depending in the first choice, see Supplementary materials).

In the Ultimatum Game, participants responded to six proposals of another anonymous participant in the study on how to divide 10 points. In the case of a rejection both players earn zero, whereas if the participant accepted the offer, the players get the proposed outcome. The first proposal was an equal split but every next proposal was more beneficial for the other than for self (i.e., (5, 5), (4, 6), (3, 7), (2, 8), (1, 9), (0, 10). For both games, we were interested in the point at which a participant switched their preference from an equal to unequal distribution, or vice versa. This allowed us to infer the point at which participants were indifferent between either distribution. This 'indifference point' represents participants' inequality aversion. That is, higher indifference points in the UG indicate stronger disadvantageous inequality aversion [range 0 - 5], whereas lower indifference points in the DG indicate stronger advantageous inequality aversion [range 0 - 10]. We used indifference points as measures of inequality aversion in all behavioral analyses.

Note that for using social preferences in the reinforcement learning models we transformed the indifference points to measures of advantageous ($\beta$) and disadvantageous ($\alpha$) inequality aversion, following the equations of (Blanco et al., 2011). Accordingly, $\alpha$ varied between 0 - 4.5 and $\beta$ varied between 0 - 1 (see Supplementary materials for a detailed description, and Figure S4). Note that these transformations are only relevant for the computational modeling as they are used for obtaining a subjective payoff matrix. However, if we rerun the behavioral analyses with these transformed inequality aversions all conclusions remain the same.

Finally, indifference points and inequality parameters $\alpha$ and $\beta$ can only be calculated for people that show consistency in choice behavior in the DG and UG. In total, 54 participants were excluded due to missing values for social preferences (missing disadvantageous inequality aversion, $n = 1$; missing advantageous inequality aversion, $n = 53$). See Supplementary materials, and Figures S2-S4 for a more detailed description of the Dictator game and Ultimatum Game, and calculation of indifference points and inequality aversion measures (α, β).

### Prior expectations

Before the start of each of the economic games (Trust Game and Coordination Game), we assessed participants' prior expectations about the behavior of other people. We asked participants "Suppose that there are 10 other players, how many of these 10 do you think will choose X?" (i.e., 'trustworthy' choice in the Trust Game, or choice to have an advantage over another person in the Coordination Game). This resulted in a prior expectation of the trustworthiness of others (Figure 3C) and a prior expectation of others' tendencies to accept an advantage (Figure 3D), both varying from 0 to 10.



### Procedure

All tests were administered in school settings. In the instruction of each learning task, three control questions were included to ensure understanding of the experimental procedure. Two questions quizzed the participant on their understanding of the point distribution (e.g., type how many points each player was winning in a certain choice combination), and one question referred to the color denotation of the two environments. If participants failed one of the control questions, the instruction was repeated until participants understood the procedure of the game. For participants younger than 12, instructions were read out load by an experimenter. All participants completed the tasks by themselves on computers in a quiet environment at school or at the university. Background variables such as the Raven SPM (estimated IQ) and several questionnaires (not relevant to the current study) were administered online using Qualtrics (www.qualtrics.com). In a separate session the DG, UG, and learning tasks were completed using the online software LIONESS Lab (Giamattei et al., 2020).

### Statistical analyses of behavioral data

To assess age-related changes in prior expectations and social preferences we ran separate robust linear regression analyses (5000 bootstraps), each with age linear and age quadratic as predictors. Multiple mediation analyses were conducted in SPSS using the computational tool PROCESS version 3.3(Hayes, 2017). For indirect effects, 95% (two-tailed) bias-corrected bootstrapped confidence intervals were calculated using 5000 repetitions. An indirect effect is significant if the confidence interval for the indirect effect does not include zero. These analyses were conducted in SPSS 25, and all tests were two-sided.

#### *Generalized linear mixed models*

To analyze choice behavior in the Trust Game and Coordination Game, we fitted logistic generalized linear mixed models (GLMMs) to decisions to choose A (coded as 0) or B (coded as 1) to each game separately. Analyses were conducted in R 3.6.1 (R Core Team, 2020), using the lme4 package (Bates et al., 2014). In all models, participant ID entered the regression as a random intercept to handle the repeated nature of the data. Where appropriate, environment was entered as a random slope in our analyses to handle the differences between individuals in their responsiveness to learning to different levels of (non)cooperation. Our GLMMs included a main effect of environment (e.g., Trustworthy environment, Untrustworthy environment), age in years (linear and quadratic), prior expectations of others' choices and social preferences, and all two-way interactions with environment (see Tables S1-S4 for all GLMM results). Note that for the Trust Game we only added disadvantageous inequality aversion, whereas for the Coordination Game both social preferences were included. That is, in the Coordination Game both types of inequality can occur and drive choice behavior, in contrast to the Trust Game in which only disadvantageous inequality is present.

In all GLMMs, age, prior expectations, disadvantageous and advantageous inequality aversion (mean-centered and scaled) and categorical predictor variables were specified by a sum-to-zero contrast (e.g., sex: -1 = boy, 1 = girl). For the mixed-effects model analyses the optimizer "bobyqa" (Powell, 2009) was used, with a maximum number of $1 \text{x} 10^5$ iterations. P-values for all individual terms were determined by Loglikelihood Ratio Tests as implemented in the mixed function in the afex package (Singmann et al., 2020). All statistics, including odds ratios and confidence intervals, are reported in Tables S1 –S6.

### *Computational modeling*

To gain a mechanistic understanding of participants' learning to adjust in the Trust Game and the Coordination Game, we used a basic reinforcement learning (RL) model(Sutton & Barto, 1998) and extended it to accommodate social preferences(Fehr & Schmidt, 1999) (aversion to unequal outcomes). All our models follow the basic logic of RL, in which agents learn about others behavior by updating their expectations with experience. In the case of the games, these expectations (denoted $p$) concern the behavior of their interaction partners (X or Y; cf. Figure 1). In each trial, $p$ is updated with a magnitude proportional to the prediction error (PE; the difference between the actual and expected choice) and the learning rate $\lambda$. Formally, $p_{t+1} = p_t + \lambda \cdot PE$, where $PE = p$ − choice of other (1 if $X$, 0 otherwise). We fit a set of reinforcement learning models to the data to investigate how $\lambda$ changes across age cohorts. This parameter is bounded between 0 (which means no updating of expectations at all) and 1 (which means that expectations match the decision of the most recent player).

In our models, the value of $p$ determines the relative weights of $w_A$ and $w_B$ (Figure 4). Each of the games is characterized by a payoff matrix (Figure 1). In each trial $t$, expected monetary payoffs of choosing $A$ or $B$ are respectively given by $w_{A,t} = p_t \cdot a + (1 - p_t) \cdot b$ and $w_{B,t} = p_t \cdot c + (1 - p_t) \cdot d$. We set the initial value of $p_0$ to the cohort mean prior measured in our experiment. The probability that a participant chooses $A$ is determined by a standard softmax function:

$\Pr(A) = [1 + e^{-\theta \cdot (w_A - w_B)}]^{-1}$ As there are only two options (A and B) to choose from, the probability of choosing *B* is simply 1 – Pr (*A*). In the softmax formula, $\theta$ reflects 'decision sensitivity' and accounts for stochasticity in participants' choices: low values of $\theta$ indicate high levels of stochasticity (Pr (*A*) and Pr (*B*) tend to be near 0.5), and high values of $\theta$ indicate low levels of stochasticity. In our model fits, $\theta$ is a free parameter allowed to vary between 0 and 5.

We extended this baseline model with two factors. First, we include the cohort mean measures of social preferences; that is, we add the measured cohort averages of *disadvantageous* and *advantageous inequality aversion* to calculate $w_A$ and $w_B$. In particular, for the Trust Game, the weight of option *A* was penalized with a value proportional to the disadvantageous inequality aversion (i.e., $\alpha$; note that we drop the subscripts as we assume social preferences to be parameters with a constant value (Fehr & Schmidt, 1999): $w_A = p \cdot a + (1 - p_t) \cdot [\, b - \alpha \cdot (b' - b) \,]$. As for option *B* the payoffs for both partners are always equal, $w_B$ is unaffected by social preferences. For the Coordination Game, social preferences can affect the weights of both *A* and *B*: $w_A = p \cdot \alpha \cdot (a' - a)$, and $w_B = p \cdot \beta \cdot (d - d')$, where $\beta$ denotes advantageous inequality aversion.

2

Second, we allowed the learning rate $\lambda$ to decay over the course of interactions. We implemented this by defining $\lambda_t = \lambda_0 \cdot r^{-\tau}$, where *r* denotes the trial number, and $\tau$ is a free parameter that reflects the speed of the decay in learning, allowed to vary between 0 and 5. The values of the estimated parameters ($\theta$, $\lambda$, $\tau$) per age cohort and per game can be found in Table S7. For each of the four age cohorts, we pooled the data and fitted the model with each possible combination of the factors 'social preferences' and 'decay', yielding a total of four models per cohort per game. Note that we also evaluated a potential role for prior expectations by including mean cohort-level prior expectations in the initial valuation of the choice options. However, because prior expectations were relatively close to 5 (range 0 – 10; Figure 2) this was close to the default expectation *p* of 0.5, marking indifference between the environments at the first choice. Hence, we did not apply formal tests of improved model fit for prior expectations.

Figure 4D shows the goodness-of-fit for each model summed across the four age cohorts relative to the best model, which includes both social preferences and decay. We included a simulation study with a parameter recovery component in the Supplementary materials. Our approach of fitting reinforcement learning models to cohort-level data was motivated by the fact that we had a limited number of observations to accurately fit our model to individual-level choice data. Note that sensitivity analyses with individually-derived parameters indicated this did not influence any of our model-fit conclusions or main findings.

## Data and code availability

The data that support the findings of this study, and all relevant R codes can be found on OSF (https://osf.io/z84g6/) and in the Leiden Repository (https://doi.org/10.34894/Z1OYYA).

## Supplementary materials

### Prior expectations and social preferences: IQ, and sex effects

In a series of robust regression analyses (5000 bootstraps) we examined the effect of IQ and sex effects on participants' social preferences (indifference points) and prior expectations across all ages. We found that estimated IQ was not related to advantageous inequality aversion, prior expectations of others' trustworthiness, and prior expectations of others' tendency to choose more for oneself (all *P*s > 0.1). However, participants with higher IQ were less disadvantageous inequality averse (IQ, B = -0.038, β = -0.308, $p < .001$, 95% CI [-0.053, -0.023]). Also, compared to boys, girls scored higher on advantageous inequality aversion (Sex, B = -1.650, β = 0.245, $p = .001$, 95% CI [-2.569, -0.730]). Compared to boys, girls also expected that others would choose to have more than the participant (Sex, B = 0.457, β = 0.135, $p = .035$, 95% CI [0.033 - 0.881]). We observed no sex differences in disadvantageous inequality aversion and prior expectations of others' trust behavior (all *p*s > .05).

Next, we assessed correlations between advantageous and disadvantageous inequality aversion, and between the prior expectations of the Trust Game and Coordination Game. Inequality aversions were slightly correlated, indicating that people who disliked being behind also liked being ahead ($r = 0.176$, $p = .012$; $N = 202$). This relation was significant even when controlled for age ($r = 0.151$, $p = .032$). Prior expectations about others in the Trust Game and the Coordination Game were not correlated ($r = 0.069$, $p = .282$; $N = 245$), neither when controlled for age ($r = 0.067$, $p = .297$).

### Developmental changes in the non-social learning task

To assess whether people are able to adjust their decision-making in a non-social context, we included a non-social learning task with two environments of computer opponents (payoff matrix $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$, see Figure S1A). Similar to the economic games, participants could maximize their payoffs by coordinating their choices to the choices of the computer opponent. That is, choosing option A when playing against the environment that most often (11/15 trials) chooses X, and choosing option B when playing against the environment that most often (11/15 trials) chooses Y.

We fitted a logistic GLMM to decisions (coded as correct/incorrect) in the non-social learning task, with age linear and age quadratic as predictors. Performance increased linearly with age (Figure S1B; see Table S5 for all statistics).

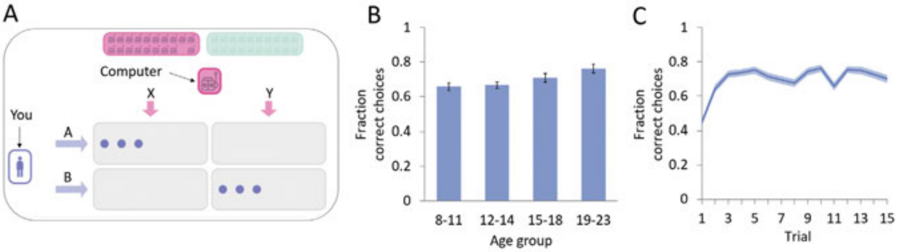



**Figure S1.** Non-social learning task. **(A)** layout of the learning task. **(B)** performance per age cohort. **(C)** performance over trials pooled across all participants (N=245). Error bars in panel b and shaded areas in panel c show standard error of the mean (s.e.m).



## Age-related differences in the non-social vs. social learning game

The main focus of this study was to compare the influence of prior beliefs, social preferences, and feedback-based updating across age groups in different social learning environments. However, the inclusion of a non-social task lends itself for the opportunity to compare learning performance in a social and non-social task more generally. Exploratively, we therefore compared performance in the non-social learning task versus the Trust Game and the non-social task versus the Coordination Game.

We fitted a binomial GLMM to decisions (coded as correct/incorrect) with age linear * type of task (social vs non-social) as predictors. Participant was included as a random intercept to account for the repeated nature of the choice data. Overall performance (grouped across environments) was lower in the Trust Game compared to the non-social learning task; main effect Task $p < 0.001$; Trust Game *Mean* accuracy = 0.613 (*SD* = 0.487); Non-social *Mean* accuracy = 0.699 (*SD* = 0.459). Follow-up tests, showed that this advantage for learning in the non-social task was present for each age group (8-11 years $p < 0.001$; 12-14 years $p < 0.001$; 15-18 years $p = .024$; 19-23 years $p < 0.001$). A similar GLMM showed that overall performance (grouped across environments) was also lower in the Coordination Game compared to the non-social task; main effect Task $p < 0.001$; Coordination Game *Mean* accuracy = 0.668 (*SD* = 0.470). Follow-up tests showed that this was, however, only the case for the two youngest age groups, and from approximately age 15 adolescents performed -on average- equally well in both games (8-11 years, $p < .001$; 12-14 years, $p < .001$; 15-18 years, $p = .870$; 19-23 years, $p = .830$). These comparisons between games confirm that specifically young adolescents find it difficult to adjust behavior in social compared to non-social games. Our analyses described in the main paper shed light on what factors underlie age-related change in learning in social environments that differ in their level of cooperation.

## Testing confounding effects of sex and IQ on choice behavior in the social-learning games

To assess possible effects of sex and IQ on choice behavior in the learning tasks, we ran additional GLMMs (one per game) that included all predictor variables from the main model (see *General linear mixed-effects models* in main text), and added main effects of sex and estimated IQ. In the Trust Game there were no main effects of sex or IQ (see SI table 2 for all statistics); all effects remained significant after adding sex and IQ to the model – except for the main effect of disadvantageous inequality aversion. In the Coordination Game, there were also no main effects of sex and IQ (see SI table 4 for all statistics). All effects remained significant after adding sex and IQ to the model. In the Non-social learning task, there was no main effect of sex (see SI table 6 for all statistics). However, there was a main effect of IQ ($p < .001$) indicating that participants with a higher IQ had better performance in the non-social learning task. These results suggest that IQ and sex differences did not influence performance on the social learning tasks, and did not influence any of the observed age-related changes in adjustment behavior. However, IQ did relate positively to a general learning tendency in the non-social learning task.

## Calculation of social preferences Dictator Game

The Dictator Game (DG) was used to estimate people's advantageous inequality aversion. The aim of the DG is to identify the point at which a participant is indifferent between (10, 0) and the equal distribution. To this end, we always had participants initially chose in the first trial between (10, 0) and (5, 5). The number of points in the subsequent trials depended on the choice in that first trial. That is, if in the first trial a participant selected the equal distribution (5, 5) rather than the (10, 0), the number of points in the equal distribution decreased with one point in each subsequent trial (i.e., (5, 5); (4, 4); (3, 3); (2, 2); (1, 1); (0, 0)). In contrast, if in the first trial the participant selected the unequal distribution (10, 0) over the equal distribution (5, 5), in the subsequent trials the number of points in the equal distribution increased to (10, 10) (i.e., (5, 5); (6, 6); (7, 7); (8, 8); (9, 9); (10, 10)). A participant's inequality aversion is determined by the equal distribution (x, x) which he/she regards as good as the (unequal) distribution (10, 0). For participants with consistent preferences (YY%), this 'indifference point' (*IP*) follows from the point at which they switch from choosing the equal distribution over the unequal distribution (or vice versa; (Blanco et al., 2011). For example, if a participant prefers (10, 0) over (6, 6), but prefers (7, 7) over (10, 0), we assume that $IP = 6.5$.

For participants who switched multiple times from equal to unequal distribution (i.e., who did not show a consistent choice pattern), we fitted a softmax function (see below) to their DG choices to approximate their indifference point. In particular, we coded a choice for the unequal distribution as 0 and the equal distribution ($x$, $x$) as 1. To these choices, 6 for each participant, we then fitted a softmax function $y = [1+\exp(-Z)]^{-1}$, where $Z = a + b*x$.

Subsequently, we solved for y = 0.5 (i.e., *-a / b*) if the fitted value of *b* was positive (i.e., when participants tended to choose equal outcomes more often when the value of *x* increased). If the approximated indifference point was outside of the theoretically possible range of [0, 10], we assumed it was undefined. This procedure resulted in approximated switch points *IP'*.

In the Dictator Game, 148 participants showed a consistent choice pattern (i.e., up to 1 switching point, 60%). For 54 participants with multiple switching points we successfully used the softmax function to approximate their indifference point. For 43 participants the approximated indifference point was outside of the theoretically possible range [0, 10], or had a negative values of *b*. These participants where therefore excluded from analyses that used the advantageous inequality aversion measure.



These calculated IP and estimated IP' were used in our behavioral analyses. To include social preferences in subsequent reinforcement learning models we transformed individuals' IPs following Blanco et al. (2011) as: β = 1 – *IP*/10. We used the general constraints as formalized by Fehr & Schmidt (1999), with the range of advantageous inequality aversion values limited between 0 and 1. We used these transformed inequality aversions (α and β) to calculate the subjective pay-off matrix in our RL modeling (see *Methods, section Computational modeling* in the main text).

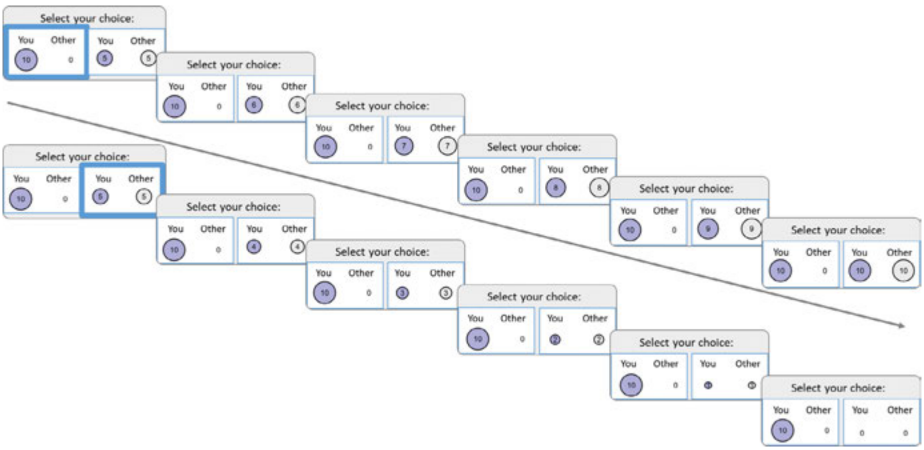



**Figure S2.** Display of the Dictator Game trials. Trial sequence depends on the choice in the first trial: upper sequence when first choice is for the unequal distribution, lower sequence when first choice is for the unequal distribution.

## Calculation of social preferences Ultimatum Game

The Ultimatum Game (UG) was used to estimate people's disadvantageous inequality aversion. The UG is a sequential two-stage game, in which a proposal for a division of points is offered by the proposer, and can be rejected or accepted by the responder. In the case of a rejection both players earn zero. If the responder accepts, players get the outcome proposed. In the proposer stage participants were presented with 10 points and could make an offer to the responder keeping 0 up to 10 points for themselves (i.e., 10 (self), 0 (other); (10, 0); (9, 1); (8, 2); (7, 3); (6, 4); (5, 5); (4, 6); (3, 7); (2, 8); (1, 0)).

In the responder stage, participants were told they were now paired with a new player and that they could accept or reject their proposals (Figure S3). As a responder, participants responded to 6 proposals. The first proposal was an equal split but every next proposal was more beneficial for the other than for self (i.e., (5, 5), (4, 6), (3, 7), (2, 8), (1, 9), (0, 10)). This responder stage was used to obtain a measure of disadvantageous inequality aversion, given that the minimum-acceptable offers in the UG is used as a point estimate of disadvantageous inequality aversion for each individual.

In the Ultimatum Game, the majority of participants showed a consistent choice pattern (i.e., up to 1 switching point, $n$ = 234, 95.5%). In total, 11 participants had multiple switching points (up to 3), and again we used the softmax function described above for the Dictator Game to approximate their indifference point. For 1 participant the approximated indifference point was outside of the theoretically possible range [0, 5]. This participant was therefore excluded from analyses that used the disadvantageous inequality aversion measure.

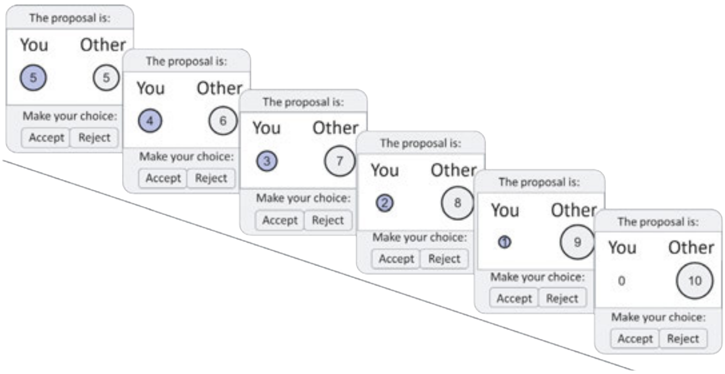



**Figure S3.** Display of the Ultimatum Game trials (responder stage).

---

Like in the DG, we used the calculated IPs and estimated IPs in our behavioral analyses. To include individuals' disadvantageous inequality aversion in subsequent reinforcement learning models, we transformed the IPs from the UG following Blanco et al. (2011): $\alpha = IP/(2*(5-IP)$. If there was no rejected offer, we set disadvantageous inequality aversion to 0. If all offers were rejected, disadvantageous inequality aversion was set to 4.5. We used these transformed inequality aversions (α and β) to calculate the subjective pay-off matrix in our RL modeling (see *Methods, section Computational modeling* in the main text).

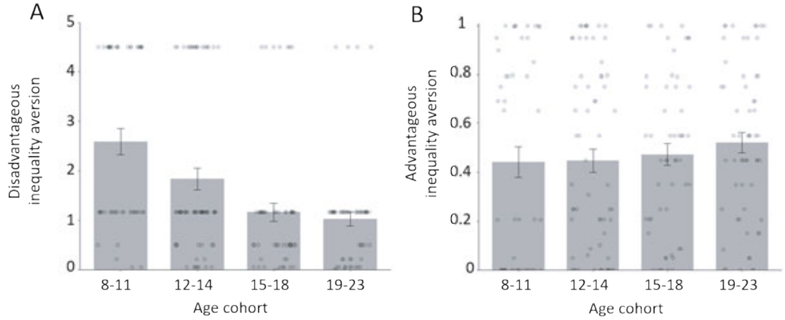



**Figure S4.** Transformed inequality aversions (α and β) per age cohort. **(A)** Transformed disadvantageous inequality aversion (α), and **(B)** transformed advantageous inequality aversion (β) which are used in the computational RL models.

2

## Computational models assessing behavioral adjustment in the non-social learning task

To assess behavioral adjustment outside of social interactions, we fitted the reinforcement models to decisions in the non-social setting (see Figure S5). Following the main text, we pool all data from participants per cohort and consider models with and without decay in learning rates across trials. As there are no other people involved, we do not consider models with or without social preferences, and assume that participants' priors in the first trial are such that they expect their computerized opponent to choose either action with equal (50%) probability.

We observe that models including a decay in learning rates lead to an improved fit. Figure S5 illustrates the estimated learning parameters of the best-fitting model. We observe that for each age cohort, learning rates decreased strongly over time. Interestingly, and in contrast to the Trust Game and the Coordination Game, the least strong decay was observed for the oldest age cohort.

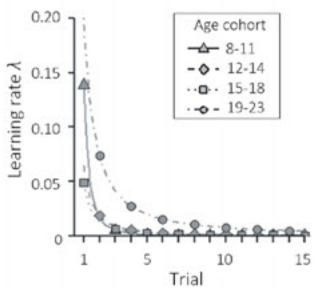



**Figure S5.** Estimated learning rates per age cohort from the best models for the Non-Social learning task. Best models include decaying learning rates. Estimates learning rates as a function of the trial number, for each of the four age cohorts separately for the Non-Social learning task.

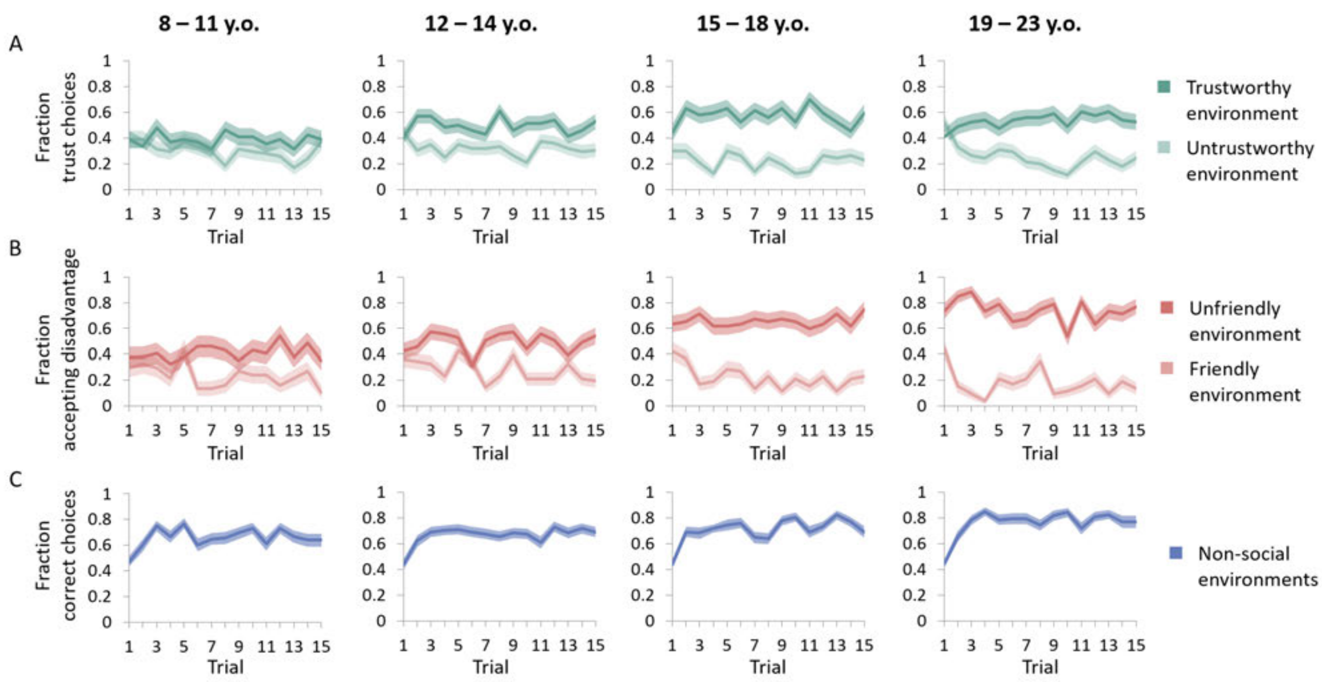



**Figure S6.** Display of choice behavior over trials per game, separately for each age cohort. Display of choice behavior over trials separately for each age cohort, in **(A)** the Trust Game ($N$ = 244), **(B)** the Coordination Game ($N$ = 202), and **(C)** the Non-social learning task ($N$ = 245). Shaded areas represent s.e.m

## Recovery of computational models

To assess the robustness of the computational models presented in the main text, we performed a recovery analysis. For each of the social games (Trust Game and Coordination Game), and for each of the age cohorts separately, we used the parameters of the best fitting models (which included social preferences and a decaying learning rate) to simulate 100 mock data sets. The distribution of mock participants across cohorts was the same as in our behavioral data set. Subsequently, we fitted each of the four models (including and excluding social preferences and decaying learning rates) to the simulated data and compared their fits. This allowed us to examine the recoverability of our models. For both economic games, for each of the 100 simulated mock data sets, the model including social preferences and decaying learning rates fitted the data best, showing that our best fitting model is recoverable.



**Supplementary Table 1. Mixed-effects model for the Trust Game (N=244)** Results of the binomial generalized linear mixed model (GLMM) testing effects of age (linear and quadratic), prior expectations, and disadvantageous inequality aversion (IA), and all 2-way interactions with environment on choice behavior in the Trust Game. Significant effects are in bold. GLMMs are described fully in the main text. Full R-code of this model is: mixed(Choice~cAge-linear*Environment + cAge-quadratic*Environment+ cDisadvantageousIA*Environment + cpriorTrustGame*Environment + (Environment | Subject), method="LRT", family=binomial, data = dat, na.action=na.exclude, control=glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000)))

| | Estimate | *SE* | $\chi^2$ | *P* | *Odds Ratio* | *CI* |
|---|---|---|---|---|---|---|
| **Environment** | -0,658 | 0,079 | 60,751 | **<0.001** | 0.52 | 0.44 – 0.60 |
| Age linear | -0,076 | 0,090 | 0,700 | 0.398 | 0.93 | 0.78 – 1.11 |
| Age quadratic | 0,112 | 0,087 | 1,633 | 0.198 | 1.12 | 0.94 – 1.33 |
| **Disadvantageous IA** | 0,215 | 0,086 | 6,102 | **0.012** | 1.24 | 1.05 – 1.47 |
| **Prior expectations** | -0,164 | 0,083 | 3,827 | **0.048** | 0.85 | 0.72 – 1.00 |
| **Environment * Age linear** | -0,307 | 0,087 | 12,203 | **<0.001** | 0.74 | 0.62 – 0.87 |
| **Environment * Age quadratic** | 0,205 | 0,084 | 5,812 | **0.015** | 1.23 | 1.04 – 1.45 |
| Environment * Disadvantageous IA | 0,148 | 0,083 | 3,139 | 0.073 | 1.16 | 0.99 – 1.36 |
| Environment * Prior expectations | 0,106 | 0,080 | 1,744 | 0.183 | 1.11 | 0.95 – 1.30 |
| Post-hoc \| *Trustworthy environment* | | | | | | |
| **Age linear** | -0,384 | 0,140 | 7,324 | **0.006** | 0.68 | 0.52 – 0.90 |
| **Age quadratic** | 0,316 | 0,136 | 5,330 | **0.020** | 1.37 | 1.05 – 1.79 |
| **Disadvantageous IA** | 0,362 | 0,134 | 7,182 | **0.007** | 1.44 | 1.11 – 1.87 |
| Prior expectations | -0,057 | 0,128 | 0,197 | 0.655 | 0.94 | 0.73 – 1.21 |

**Supplementary Table 1.** *Continued*

| | Estimate | *SE* | $\chi^2$ | *P* | *Odds Ratio* | *CI* |
|---|---|---|---|---|---|---|
| Post-hoc \| *Untrustworthy environment* | | | | | | |
| **Age linear** | 0,233 | 0,108 | 4,618 | **0.031** | 1.26 | 1.02 – 1.56 |
| Age quadratic | -0,093 | 0,104 | 0,785 | 0.372 | 0.91 | 0.74 – 1.12 |
| Disadvantageous IA | 0,068 | 0,103 | 0,428 | 0.510 | 1.07 | 0.88 – 1.31 |
| **Prior expectations** | -0,270 | 0,101 | 7,125 | **0.007** | 0.76 | 0.63 – 0.93 |

**Supplementary Table 2. Mixed-effects model with IQ and sex effects for the Trust Game (N=244)**
Results of the binomial generalized linear mixed model (GLMM) testing confounding effects of IQ and sex on choice behavior in the Trust Game. Predictors of interest are underlined. GLMMs are described fully in the main text. Full R-code of this model is: mixed(Choice~IQ + sex + cAge-linear*Environment + cAge-quadratic*Environment+ cDisadvantageousIA*Environment + cpriorTrustGame*Environment + (Environment | Subject), method="LRT", family=binomial, data = dat, na.action=na.exclude, control=glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000)))

| | Estimate | *SE* | $\chi^2$ | *P* | *Odds Ratio* | *CI* |
|---|---|---|---|---|---|---|
| <u>Sex</u> | 0.047 | 0.084 | 0.316 | 0.574 | 1.05 | 0.89 – 1.24 |
| <u>IQ</u> | 0.107 | 0.087 | 1.501 | 0.216 | 1.11 | 0.94 – 1.32 |
| Environment | -0.656 | 0.079 | 60.201 | <0.001 | 0.52 | 0.44 – 0.61 |
| Age linear | -0.070 | 0.093 | 0.556 | 0.451 | 0.93 | 0.78 – 1.12 |
| Age quadratic | 0.102 | 0.088 | 1.317 | 0.248 | 1.11 | 0.93 – 1.32 |
| Disadvantageous IA | 0.244 | 0.091 | 7.103 | 0.007 | 1.28 | 1.07 – 1.52 |
| Prior expectations | -0.156 | 0.084 | 3.419 | 0.062 | 0.86 | 0.73 – 1.01 |
| Environment * Age linear | -0.307 | 0.087 | 12.189 | <0.001 | 0.74 | 0.62 – 0.87 |
| Environment * Age quadratic | 0.205 | 0.084 | 5.806 | 0.015 | 1.23 | 1.04 – 1.45 |
| Environment * Disadvantageous IA | 0.149 | 0.083 | 3.163 | 0.072 | 1.16 | 0.99 – 1.36 |
| Environment * Prior expectations | 0.107 | 0.080 | 1.754 | 0.182 | 1.11 | 0.95 – 1.30 |

**Supplementary Table 3. Mixed-effects model for the Coordination Game (N=202)** Results of the binomial generalized linear mixed model (GLMM) testing effects of age (linear and quadratic), prior expectations, and disadvantageous inequality aversion (IA), and all 2-way interactions with environment on choice behavior in the Coordination Game. Significant effects are in bold. GLMMs are described fully in the main text. Full R-code of this model is: mixed(Choice~cAge-linear*Environment + cAge-quadratic*Environment+ cDisadvantageousIA*Environment + cAdvantageousIA*Environment + cpriorCoordinationGame*Environment + (Environment | Subject), method=”LRT”, family=binomial, data = dat, na.action=na.exclude, control=glmerControl(optimizer = “bobyqa”, optCtrl = list(maxfun = 100000)))

| | Estimate | *SE* | $\chi^2$ | *P* | *Odds Ratio* | *CI* |
|---|---|---|---|---|---|---|
| **Environment** | -1,041 | 0,086 | 115,264 | **<0.001** | 0.35 | 0.30 – 0.42 |
| **Age linear** | -0,239 | 0,071 | 11,051 | **0.001** | 0.79 | 0.69 – 0.90 |
| Age quadratic | 0,124 | 0,068 | 3,213 | 0.070 | 1.13 | 0.99 – 1.29 |
| **Disadvantageous IA** | 0,134 | 0,068 | 3,821 | **0.048** | 1.14 | 1.00 – 1.31 |
| **Advantageous IA** | 0,172 | 0,066 | 6,689 | **0.009** | 1.19 | 1.04 – 1.35 |
| Prior expectations | 0,035 | 0,065 | 0,281 | 0.594 | 1.04 | 0.91 – 1.18 |
| **Environment * Age linear** | -0,458 | 0,093 | 23,002 | **<0.001** | 0.63 | 0.53 – 0.76 |
| Environment * Age quadratic | 0,067 | 0,090 | 0,537 | 0.460 | 1.07 | 0.90 – 1.28 |
| Environment * Disadvantageous IA | 0,068 | 0,090 | 0,570 | 0.450 | 1.07 | 0.90 – 1.28 |
| **Environment * Advantageous IA** | 0,260 | 0,087 | 8,746 | **0.003** | 1.30 | 1.09 – 1.54 |
| Environment * Prior expectations | 0,030 | 0,086 | 0,117 | 0.731 | 1.03 | 0.87 – 1.22 |
| Post-hoc \| *Friendly environment* | | | | | | |
| **Age linear** | 0,209 | 0,100 | 4,266 | **0.037** | 1.23 | 1.01 – 1.50 |
| Age quadratic | 0,058 | 0,097 | 0,351 | 0.551 | 1.06 | 0.88 – 1.28 |
| Disadvantageous IA | 0,077 | 0,095 | 0,637 | 0.420 | 1.08 | 0.90 – 1.30 |
| Advantageous IA | -0,083 | 0,093 | 0,787 | 0.372 | 0.92 | 0.77 – 1.10 |
| Prior expectations | 0,011 | 0,092 | 0,013 | 0.909 | 1.01 | 0.84 – 1.21 |
| Post-hoc \| *Unfriendly environment* | | | | | | |
| **Age linear** | -0,695 | 0,130 | 26,904 | **<0.001** | 0.50 | 0.39 – 0.64 |
| Age quadratic | 0,193 | 0,126 | 2,308 | 0.126 | 1.21 | 0.95 – 1.55 |
| Disadvantageous IA | 0,192 | 0,126 | 2,289 | 0.128 | 1.21 | 0.95 – 1.55 |
| **Advantageous IA** | 0,425 | 0,122 | 11,718 | **<0.001** | 1.53 | 1.20 – 1.94 |
| Prior expectations | 0,064 | 0,121 | 0,279 | 0.595 | 1.07 | 0.84 – 1.35 |

**Supplementary Table 4. Mixed-effects model with IQ and sex effects for the Coordination Game (N=202)** Results of the binomial generalized linear mixed model (GLMM) testing confounding effects of IQ and sex on choice behavior in the Coordination Game. Predictors of interest are underlined. GLMMs are described fully in the main text. Full R-code of this model is: mixed(Choice~IQ + sex + cAge-linear*Environment + cAge-quadratic*Environment+ cDisadvantageousIA*Environment + cAdvantageousIA*Environment + cpriorCoordinationGame*Environment + (Environment | Subject), method="LRT", family=binomial, data = dat, na.action=na.exclude, control=glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000)))

| | Estimate | *SE* | $\chi^2$ | *P* | *Odds Ratio* | *CI* |
|---|---|---|---|---|---|---|
| <u>Sex</u> | -0,012 | 0,067 | 0,031 | 0.860 | 0.99 | 0.87 – 1.13 |
| <u>IQ</u> | -0,104 | 0,068 | 2,286 | 0.127 | 0.90 | 0.79 – 1.03 |
| Environment | -1,043 | 0,086 | 115,527 | <0.001 | 0.35 | 0.30 – 0.42 |
| Age linear | -0,226 | 0,071 | 9,758 | 0.002 | 0.80 | 0.69 – 0.92 |
| Age quadratic | 0,132 | 0,068 | 3,648 | 0.054 | 1.14 | 1.00 – 1.30 |
| Disadvantageous IA | 0,101 | 0,071 | 2,015 | 0.151 | 1.11 | 0.96 – 1.27 |
| Advantageous IA | 0,175 | 0,067 | 6,740 | 0.008 | 1.19 | 1.05 – 1.36 |
| Prior expectations | 0,023 | 0,066 | 0,126 | 0.722 | 1.02 | 0.90 – 1.16 |
| Environment * Age linear | -0,457 | 0,093 | 22,829 | <0.001 | 0.63 | 0.53 – 0.76 |
| Environment * Age quadratic | 0,066 | 0,090 | 0,527 | 0.464 | 1.07 | 0.89 – 1.28 |
| Environment * Disadvantageous IA | 0,068 | 0,090 | 0,581 | 0.445 | 1.07 | 0.90 – 1.28 |
| Environment * Advantageous IA | 0,260 | 0,087 | 8,737 | 0.003 | 1.30 | 1.09 – 1.54 |
| Environment * Prior expectations | 0,030 | 0,086 | 0,117 | 0.731 | 1.03 | 0.87 – 1.22 |

**Supplementary Table 5. Mixed-effects model for the Non-social learning task (N=245)** Results of the binomial generalized linear mixed model (GLMM) testing effects of age (linear and quadratic) on choice behavior in the Non-social Game. Significant effects are in bold.

| | Estimate | *SE* | $\chi^2$ | *P* | *Odds Ratio* | *CI* |
|---|---|---|---|---|---|---|
| **Age linear** | **0.251** | **0.082** | **9.150** | **0.002** | **1.29** | **1.09-1.51** |
| Age quadratic | 0.143 | 0.083 | 2.998 | 0.083 | 1.15 | 0.98-1.36 |

**Supplementary Table 6. Mixed-effects model with IQ and sex effects for the Non-social learning task (N=245)** Results of the binomial generalized linear mixed model (GLMM) testing confounding effects of IQ and sex on choice behavior in the Non-social learning task. Predictors of interest are underlined.

| | Estimate | *SE* | $\chi^2$ | *P* | *Odds Ratio* | *CI* |
|---|---|---|---|---|---|---|
| Age linear | 0.183 | 0.082 | 4.936 | 0.026 | 1.20 | 1.02-1.41 |
| Age quadratic | 0.114 | 0.080 | 2.027 | 0.155 | 1.12 | 0.96 -1.31 |
| <u>IQ</u> | 0.334 | 0.075 | 18.963 | <.001 | 1.40 | 1.20-1.62 |
| <u>Sex</u> | -0.074 | 0.078 | 0.892 | 0.345 | 0.93 | 0.80-1.08 |

2

**Supplementary Table 7. Parameters of best fitting computational models per game** Table with all parameters of the best fitting computational models (i.e., models which include social preferences and decaying learning rates) per game and per age cohort. Results of the non-social game are also included.

| | | **TRUST GAME** | | |
|---|---|---|---|---|
| **Age cohort** | **prior** | **learning rate $t$=0 ($\lambda_0$)** | **decay ($\tau$)** | **decision sensitivity ($\theta$)** |
| 8-11 | 0.56 | 0.12 | 0.00 | 0.15 |
| 12-14 | 0.54 | 0.82 | 1.27 | 0.12 |
| 15-18 | 0.49 | 1.00 | 5.00 | 0.23 |
| 19-23 | 0.50 | 0.96 | 5.00 | 0.24 |
| | | **COORDINATION GAME** | | |
| **Age cohort** | **prior** | **learning rate $t$=0 ($\lambda_0$)** | **decay ($\tau$)** | **decision sensitivity ($\theta$)** |
| 8-11 | 0.71 | 0.67 | 1.94 | 0.63 |
| 12-14 | 0.67 | 0.61 | 1.31 | 0.46 |
| 15-18 | 0.65 | 1.00 | 5.00 | 0.53 |
| 19-23 | 0.68 | 1.00 | 5.00 | 0.69 |
| | | **NON-SOCIAL LEARNING TASK** | | |
| **Age cohort** | **prior** | **learning rate $t$=0 ($\lambda_0$)** | **decay ($\tau$)** | **decision sensitivity ($\theta$)** |
| 8-11 | 0.50 | 0.14 | 2.99 | 1.60 |
| 12-14 | 0.50 | 0.06 | 1.80 | 2.93 |
| 15-18 | 0.50 | 0.05 | 1.87 | 4.80 |
| 19-23 | 0.50 | 0.20 | 1.43 | 1.65 |

## Participant instructions

Here we included the instructions for the Trust Game and the non-social learning task. Original instructions were in Dutch. The instructions for the Coordination Game are highly similar to instructions for the Trust Game, and are therefore not included here. Note that on almost every introduction screen for the behavioral tasks, figures of the task were included. Here the figures are only shown when necessary for understanding the accompanying text. The instructions also included some control questions; participants could only continue to the next screen when answered correctly. Full testing and instruction materials can be obtained from the corresponding author.

### Instructions Trust Game

---

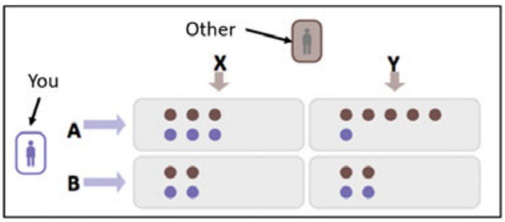
The game you will play now has 3 parts. Each part has 30 short rounds.
In each round you will make 1 choice.
On the next screens, the game will be explained.

---

In each round you play with another student from another school who also participated in this game.
In each round you will both make 1 choice.
In every round, the other is a **new** person.

---

The game will look like this:

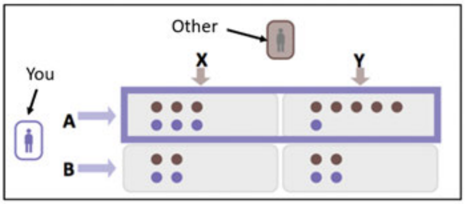



You are the **purple** icon on the left, and you can choose
between the 2 purple arrows (**A** and **B**).
The other is shown on the top of the screen, and can choose between the 2 top arrows (**X** and **Y**).

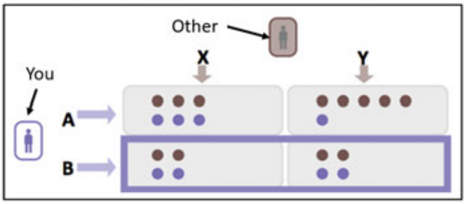
---

In each of the four boxes you see dots. Each dot represents 1 point.
The points that you can win with your choice are always **purple**.
The points that the other player can win, have the other color.

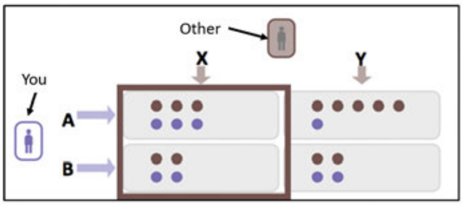
In every round you will make a choice between **A** and **B**.
The other will make a choice between **X** and **Y**.
Your choices **together** determine how many points you and the other win.



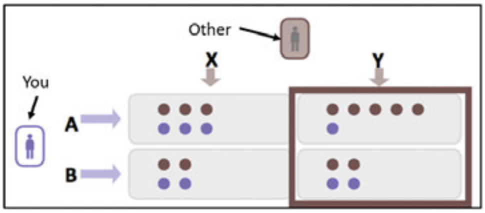

If you choose the **purple** arrow **A**, you select the two top boxes.




If you choose the **purple** arrow **B**, you select the two bottom boxes.




The other can choose arrow **X (top)** to select the two boxes on the left.




The other can choose arrow **Y (top)** to select the two boxes on the right.

---

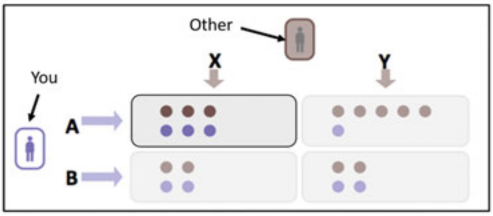
The number of points **<u>you</u>** win, depends on your choice (**A** or **B**), but <u>also</u> on the *other player*'s choice (**X** or **Y**).

The number of points **<u>the other player</u>** wins, depends on the other player's choice (**X** or **Y**), but also on *your choice* (**A** or **B**).

Only *after* you submit your choice, you can see the other's choice.

---

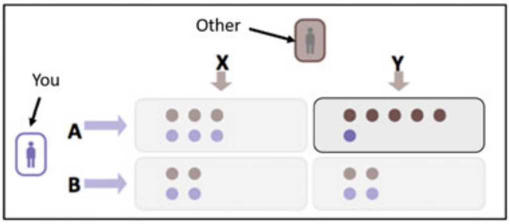
For example:




Suppose that you choose **A**, and the other chooses **X**:
- You would win **3 points**
- And the other would also win **3 points**

---

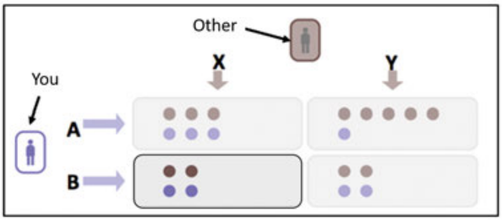
For example:




Suppose that you choose A, and the other chooses Y:
- You would win **1 point**
- And the other would win **5 points**

---

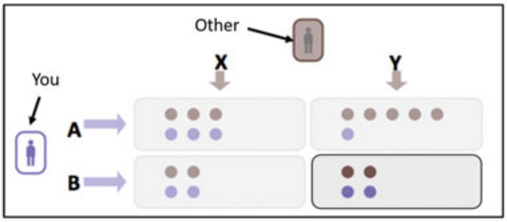
For example:




Suppose that you choose B, and the other chooses X:
- You would win **2 points**
- And the other would also win **2 points**

---

For example:




Suppose that you choose B, and the other chooses Y:
You would win ...... points (fill in)
And the other would win ...... points (fill in)

---

2

**That is correct!**

After you see what you and the other have won, the round is over and you will play with a **new person**.

**Important**: The choices of the other players are made previously by other students from another school.

---

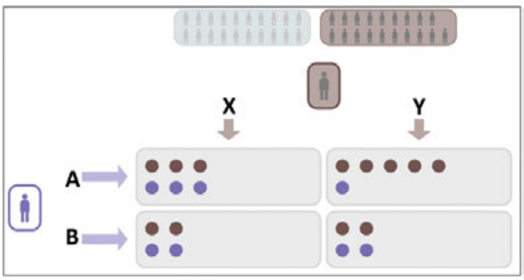
At the end of the game, the computer will select 5 random rounds.
Each point in those 5 rounds is worth 1 lottery ticket.
After all games, we will draw 1 lottery ticket from all lottery tickets at your school.
The winner receives the gift voucher!
So the more points you have, the larger your chance to win.


---

The other players can also earn lottery tickets with their choices. With the lottery tickets they also have a chance to win a gift voucher at their school.
Remember: your choices may affect the number of points **you** can win, but they may also affect the number of points the **others** can win.

---

## Environments




There are 2 environments of other players, each with their own color.
The color of the other player's icon tells you what environment they are in.

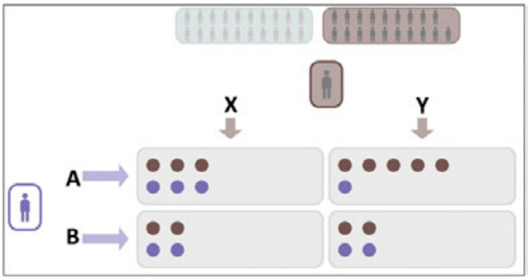
**In one environment, players usually choose X (left), and in the other environment players usually choose Y (right).**

---

In each round you will play with a new player from 1 of the 2 environments.

---

**Check Question**




Suppose that you choose **B** and the other chooses **X**.
How many points would you win? (fill in)____
How many points would the other win? (fill in)____

---

**You have correctly answered the check question!**
There will now be 2 practice rounds before you start the real game.

Make your choice by clicking on 1 of your arrows. Then click 'Confirm' at the bottom of the page.

Next, you can see the other's choice, and how many points you and the other have won in this round.
Then click 'Continue' to start with the next round.

---

*{2 practice rounds}*

---

**Check question**
Which environment did the Other player belong to?
If you don't remember, please click back.

---

The game is about to start.
Do you have any questions? Please ask the researcher now.
You can click **start** if you understand the game.

**Try to win as many points as possible by paying close attention to what the other players are doing!**

---

**End of part 1**

You are finished with part 1 of the game!

---

## Instructions Non-social learning task

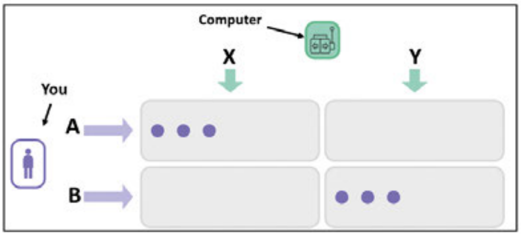
**Part 3**

In Parts 1 and 2 you were playing with other people.

In Part 3, you will not be playing with other people. Instead, you will play with computers.

---

The game will look like this:




You are shown in **purple** on the left, and you can choose between the two purple arrows (A and B).

The computer can choose between the **arrows** on the top.

Your points are the **purple** dots.

---

As before, you will choose A or B (top or bottom boxes)

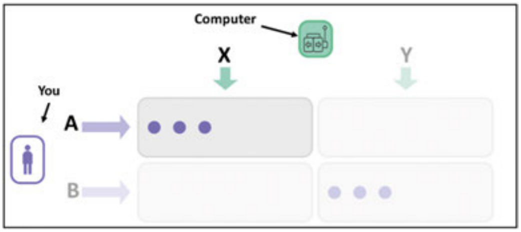
We have programmed the computer to choose X or Y (left or right boxes)

Your choice AND the computer's choice together determine how many points you will win.

**Important: the number of points you can win with your choice is different than in the previous games.**

---

For example,

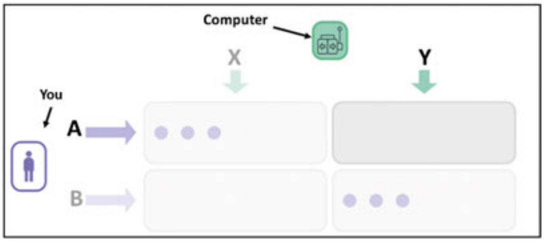



If you choose A and the computer chooses X:

You would win **3 points**

---

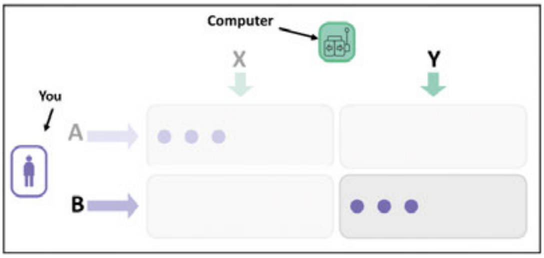
For example,




If you choose A and the computer chooses Y:
You would win **0 points**

---

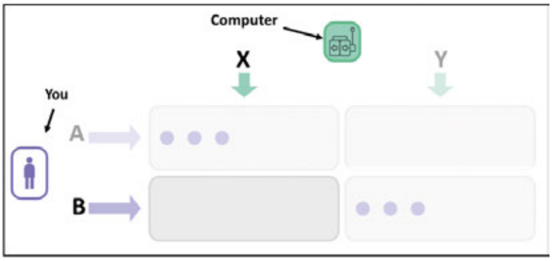
For example,




If you choose B and the computer chooses Y:
You would win **3 points**

---

For example,



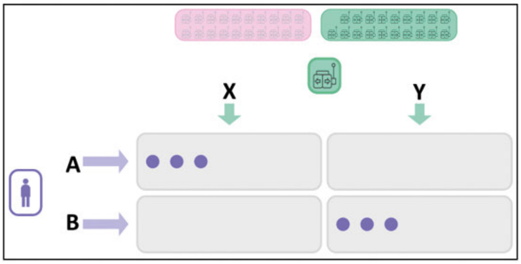

If you choose B and the computer chooses X:
You would win ____ **points** (fill in)

---

**That is correct!**
After you see what you have won, the round is over and you will play with a **new computer**.

At the end of the game, 5 rounds will be randomly selected.
Again, each point in the selected rounds is worth 1 lottery ticket.
Remember that you can win the gift voucher.
The more points you have the larger your chance of winning.

**Environments**



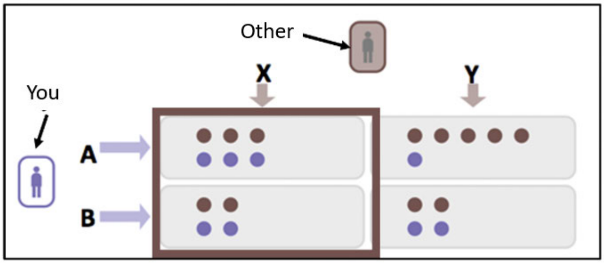

There are 2 environments of computers, each with their own color.
The computer's color tells you what environment the computer is in.
In one environment the computers usually choose **X** (left), and in the other environment the computers usually choose **Y** (right).
In each round you will play with a new computer from 1 of the 2 environments.

Click **Start** to start the game.

## Prior expectations

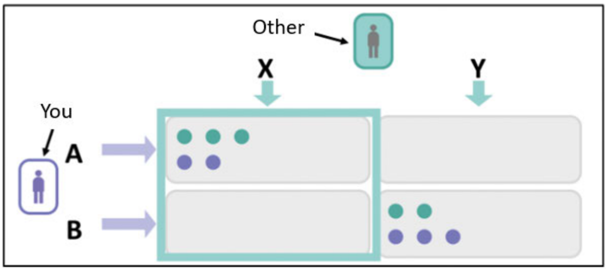
At the end of the learning task instructions, right before the start of the task, we asked about the prior expectations of the participants. This was only done for the social games (Trust Game and Coordination Game).

### *Prior expectations Trust Game*




We are curious about what you think of the other players in this game.
Suppose that there are 10 other players, how many of these 10 do you think will choose X?

### *Prior expectations Coordination Game*




We are curious about what you think of the other players in this game.
Suppose that there are 10 other players, how many of these 10 do you think will choose X?