Developmental Meritocracy *

Pablo Brañas-Garza
Universidad Loyola Andalucía

Isabelle Brocas
University of Southern California
IAST and CEPR

Juan D. Carrillo
University of Southern California
IAST and CEPR

August 2026

Abstract

We recruit 2,280 children and teenagers aged 5 to 18 in Spain to examine how young impartial observers balance merit and fairness when allocating rewards between a winner and a loser of a competition task. Using a third-party observer game, we manipulate how winners are determined, through talent, effort, or luck, and compare two reallocation methods: redistribution of tokens from winner to loser versus adding tokens to both. Participants allocate less to the loser when it requires taking tokens from the winner. Under addition, the estimated age profile rises through early adolescence and then tapers off, especially for final allocations; under redistribution, the corresponding profile is comparatively flat. Pooled estimates also indicate lower transfers to the loser when the winner is selected by talent rather than luck, while estimates for effort relative to luck are small or imprecise. Finally, allocation decisions are related to gender, personality traits, and the observer's capacity for empathy, with male participants exhibiting more meritocratic allocation patterns. Overall, our results indicate that the development of redistributive preferences is sensitive to the institutional mechanism through which redistribution is implemented.

Keywords: lab-in-the-field experiment, developmental decision making, third-party observer, inequality.
JEL Classification: C93, D63.

*We thank the members of the Los Angeles Behavioral Economics Laboratory (LABEL) and the Loyola Behavioral Lab for their insights and to Chris Crabbe for his outstanding programming skills. Special thanks to Monica Vasco, Pablo Montero and Javier Gonzalez for outstanding research assistance. We also thank the editor and referees for their constructive comments. We are grateful to the staff of Colegios Salesianos –in particular Andrés Corbacho, Manuel Redondo and Monica Soliss– for their help and support in running the experiment in their schools. The study was conducted with the University of Southern California IRB approval UP-12-00528. We gratefully acknowledge the financial support of the National Science Foundation grant SES-2315770, the Spanish Ministry of Economy and Competitiveness (PID2024-156629NB-I00), the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement number 101095175 and the UK Government, and the French National Research Agency (ANR) grant ANR-17-EURE-0010 (Investissements d'Avenir program).

1 Introduction

Redistributive preferences are a central issue in policy debates. An individual's tolerance for inequality depends on four key factors: who benefits or suffers from inequality, the perceived cause of inequality, the method of redistribution, and the attributes of the individual making the judgment. Also, redistributive preferences evolve over time, raising important questions about the interaction between innate dispositions and life experiences, particularly during childhood and adolescence.

This study focuses on *normative redistributive preferences*, examining how impartial observers decide to allocate resources between a wealthier and a poorer individual. We investigate how redistributive judgments change with age, tracing these preferences from childhood through adolescence and linking them to affective development and life experiences. Additionally, we explore how attitudes toward inequality vary depending on its source and how different redistribution mechanisms influence these preferences.

Redistributive preferences in adults have been extensively studied in the experimental literature.[1] Existing work, particularly through dictator and ultimatum games, has documented behavioral tendencies such as altruism, egalitarianism, envy, and inequity aversion when individuals have a personal stake in redistribution (Engel, 2011; Oosterbeek et al., 2004; Fehr and Schmidt, 2006). These studies also demonstrate that individuals are sensitive to the framing of allocation decisions (List, 2007) as well as to the origin of inequality (Fong, 2001; Konow, 2000). However, normative preferences, those expressed by third-party observers, have received comparatively less attention. An exception is the work of the FAIR group, which has examined how impartial adult observers allocate resources based on the source of inequality. Cappelen et al. (2022) find that observers are more likely to favor equitable distributions when disparities result from luck rather than productivity. Almås et al. (2020) show that while many Americans consider both luck- and productivity-based inequalities fair, Norwegians are more likely to reject them as unfair. These findings highlight the complexity of redistributive preferences among adults.

Our allocation-method comparison is related to, but distinct from, the give-versus-take manipulation in List (2007). The author studies self-interested adult dictators and expands their action set to include taking from the recipient. We instead study impartial observers who are children and adolescents and, most importantly, hold the feasible set of final allocations fixed across treatments. Thus, our comparison asks whether the procedure used to implement the same final allocations affects normative judgments, and how this sensitivity develops from childhood through adolescence.

Research on the development of social preferences from childhood to adulthood has focused mainly on dictator games. A pioneering study by Fehr et al. (2008) examined the

[1] Examples include Charness and Rabin (2002) on the nature of social preferences, Cappelen et al. (2013) on fairness views related to risk-taking, and Andreoni et al. (2020) on ex-ante vs. ex-post fairness, among others.

choices of children aged 3 to 8 in three binary dictator games and found that preferences for equality increase with age. Fehr et al. (2013) extended the analysis to children aged 8 to 17 and demonstrated a growing preference for (costless) egalitarianism across the sample.[2] In Almås et al. (2010), children aged 10 to 18 first completed a production task in which earnings depended on both effort and randomly assigned luck, and then decided how to distribute these earnings between themselves and a partner. The authors show that the prevalence of strict egalitarian views declines with age, whereas meritocratic views become increasingly prevalent. These studies suggest that context effects (triggering either selfish or deontological concerns) may lead to different behavioral patterns.

Distributive preferences of third-party observers in a developmental setting have rarely been studied in incentivized economic experiments. Angerer et al. (2017) and Barron et al. (2023) are two exceptions that focus on in-group favoritism based on language and nationality, respectively (see Sutter et al. (2019)'s survey).

Studying third-party observers is important not only because real-world redistribution decisions in policymaking, judicial settings, and workplaces are often made by such actors, but also because it isolates decision-makers with no material stakes, thereby removing self-interest from the equation. This allows for a cleaner understanding of how individuals of different ages justify inequality without the confounding influence of personal implication. Moreover, a systematic examination of these preferences under different conditions can provide insights into why people respond differently to inequality across contexts and, ultimately, how institutional design can influence redistributive outcomes.

For our study, we recruited a large population of 2,280 children and teenagers aged 5 to 18 years old from low- to middle-income families in Spain. Participants acted as impartial observers, deciding how to allocate resources between two players competing in a task. The study incorporates three competition tasks, each aimed at capturing different sources for inequality: *luck*, *effort* and (a novel) *talent* task. In each case, the winner receives a higher payoff than the loser. We also implement two reallocation methods. First, a take-from-the-rich method (as in Cappelen et al. (2022) and Almås et al. (2020)), where the observer can *redistribute* payoffs from the winner to the loser. Second, a novel top-up method, called *addition*, where the observer decides how to allocate additional payoffs between the winner and the loser. Importantly, the feasible set of final allocations is identical in both cases.

Impartial observers also participate as competitors. Participants first choose an allocation as outside observers, then engage in a competitive task, and finally make a second allocation decision, with the option to revise their initial choice. This sequence allows us to examine how allocations change after participants experience the task. Because all

[2] Brañas Garza (2024) replicates Fehr et al. (2008)'s experiment with 2,525 Spanish teenagers and finds that egalitarianism increases across the entire sample, while altruism drops sharply. Brocas and Carrillo (2020b) examine children aged 7 to 16 playing a dictator game where they can manipulate information to hide undesired consequences, and find that 20% to 30% do so.

participants are invited to reconsider their initial choice, the design does not separately identify the effect of experience from reflection or repeated measurement. Finally, we introduce two novel child-friendly non-cognitive tasks: an affective theory-of-mind test and a pictorial personality questionnaire.

We obtain the following insights. First, participants allocate less to the loser under redistribution than under addition, consistent with a reluctance to take rewards from the winner. Second, the estimated age trajectory is more pronounced under addition, where transfers to the loser rise through middle school and decline afterwards in the final allocation; the corresponding profile under redistribution is comparatively flat. Third, pooled estimates indicate lower transfers to the loser when the winner is selected by talent rather than by luck, whereas effort-luck estimates are small or imprecise. Fourth, revisions are systematically associated with participants' perceptions of their own performance and, less consistently, of task difficulty: individuals who believed they performed well and found the competition easy are more inclined to increase the reward of the winner. These patterns are also related to gender, personality traits, and affective processes. In particular, more empathic and agreeable individuals favor ex-post fairness, whereas more conscientious individuals, as well as male participants, display a stronger tendency toward meritocratic allocations.

Taken together, the results relate attitudes toward inequality to cognitive development, experience, and personality traits. The distinction between redistribution and addition reveals that procedural sensitivity is present among children and adolescents. Our estimates also suggest that the perceived basis of success matters: the talent-luck contrast is clearest in pooled final-allocation specifications, while the evidence for effort is less precise. Finally, post-task revisions are correlated with participants' own experience and perceptions, although our design cannot separate this channel from reconsideration of the initial choice.

More generally, the developmental patterns we estimate differ from those documented in dictator games, which suggests that redistribution by an observer and redistribution by an affected party should be studied separately. Because these patterns depend on the allocation mechanism and change with age, both developmental and contextual factors matter for attitudes toward redistribution.

The paper is organized as follows. In section 2, we present the experimental design. Section 3 develops a formal model. In section 4, we discuss the main results, including the effect on choice of age, origin of inequality and allocation mechanism. We also study the influence of personality and affective theory of mind on allocation decisions. In section 5, we present some concluding remarks.

2 Experimental design and procedures

We recruited children and adolescents to study the decision of an impartial observer who allocates rewards between two competitors. We follow the guidelines presented in Brocas and Carrillo (2020a) to address the methodological challenges that arise in studies with young participants.

Experiment. The experiment consists of two games programmed in 'oTree' (Chen et al., 2016). This study focuses on the first game, the third-party dictator. After a short break, participants played a two-person game with incomplete information, which we analyze in a separate paper (Alfonso et al., 2025).[3]

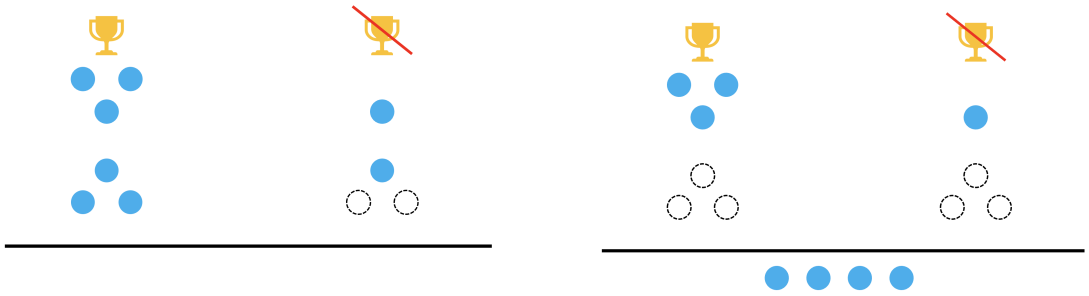
Timing. We first informed participants that they all would play an "observer task" followed by a "competition task". We emphasized that if a participant acted as the observer for two competitors, neither of those competitors would serve as their observer. In the observer task, participants allocated tokens between a future winner and a future loser of the competition task. In the competition task, they competed in pairs for a prize. We also informed them that feedback would be provided only at the end of the experiment. After the competition task, they were given the option to modify their decision in the observer task, with clear emphasis that no further modifications would be possible afterwards. These tasks were followed by the second game, an affective theory-of-mind task, and a Big 5 personality questionnaire (both described at the end of this section). The instructions for the observer and competition tasks are provided in Appendix A1.

Observer task. The impartial observer game used a visually intuitive interface (see Figure 1). Participants allocated tokens by dragging them between the winner (depicted by a trophy) and the loser (depicted by a crossed-out trophy). They could adjust the allocation freely until pressing OK to finalize their decision.

We explored two different environments for allocating tokens. First, a *redistribution* method, characterized by an initial allocation of (6,2) between the winner and the loser, with the observer able to redistribute 0, 1, or 2 tokens from the winner to the loser. This design is similar to the existing literature (Cappelen et al., 2022; Almås et al., 2020). The second is an *addition* design, where observers do not take tokens from the winner. Instead, the initial allocation is (3,1), and 4 additional tokens must be distributed between the two competitors, with at least one token allocated to each. Figures 1a and 1b show screenshots of both environments. Note that the three possible final allocations are identical in both environments: a "meritocratic" split (6,2) that maximizes the difference between winner and loser (while ensuring both receive a positive amount), a "compensatory" split (4,4) that equalizes final payoffs regardless of performance, and an "equitable" split (5,3) that

[3]The two games are distinct enough that we are not worried about cross-contamination. However, as a precautionary measure, we always conducted the two games in the same order, anonymously re-matching subjects between the two games, and did not disclose any result until the conclusion of the second game.

provides an intermediate option. However, we expect that the framing of the allocation method, take-from-the-rich (*redistribution*) vs. top-up (*addition*), may trigger different behavioral responses.



(a) *Redistribution*: 0, 1, or 2 tokens can be taken from winner and transferred to loser

(b) *Addition*: 4 tokens (bottom) must be allocated between winner and loser (with at least one for each)

Figure 1: Screenshots of the observer task: redistribution (left) v. addition (right)

Competition task. We explored three procedures for determining the winner of the competition, which encompass the three fundamental factors of performance in conventional economic models: *effort*, *talent* and *luck*.[4]

In the *effort* task, participants were randomly and anonymously paired to compete in an effort-intensive activity requiring patience and concentration. Following Gill and Prowse (2012), we employed a slider task. Each participant faced 39 non-aligned sliders, each starting at a random position between 0 and 100 (but never 50). The participant who successfully positioned the most sliders at exactly 50 within 90 seconds was the winner.[5] Prior to the competition, participants completed a short practice round where they adjusted three sliders to 50. Figure 2a provides a screenshot of the effort competition task, showing four sliders already set at 50 and a timer indicating 31 seconds remaining.

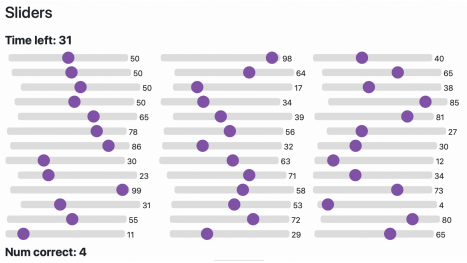
In the *talent* task, participants answered a series of 8 to 12 age- and culture-appropriate multiple-choice general knowledge questions, covering topics such as mathematics, biology, history, sports, and entertainment. Participants had, depending on age, 12 to 18 seconds to select one of four possible answers. The participant who answered the most questions correctly was declared the winner. Before the competition, participants completed a practice question to familiarize themselves with the format. Figure 2b displays a screenshot of the talent competition task (translated from the original Spanish version).

In the *luck* task, participants were simply informed that the computer would randomly

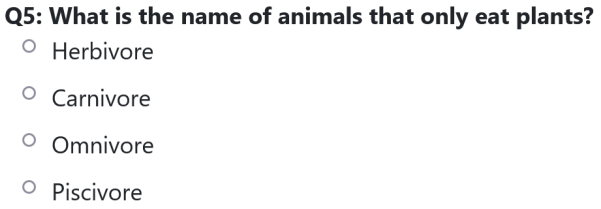
[4]Canonical models of performance are typically summarized by $y = f(e, \theta) + \eta$, where y = performance, e = effort, θ = ability, and η is a stochastic component.

[5]In case of a tie, the participant who reached that number of sliders first was the winner. Participants were never informed of the exact performance difference between winner and loser.

select a winner and a loser from each pair, with equal probability.



(a) *Effort*: participants had 90 seconds to position as many sliders as possible at 50



(b) *Talent*: participants answered age-appropriate, multiple choice general knowledge questions

Figure 2: Screenshots of the competition task: effort (left) v. talent (right)

Overall, the experiment consists of six variants, following a 3×2 between-subject design: three conditions for determining the winner (*effort*, *talent*, and *luck*) and two treatments for the allocation of tokens (*addition* and *redistribution*). Crucially, the design varies these two dimensions orthogonally, so differences in outcomes can be attributed to each dimension without confounding.

Population. We enrolled an unusually large sample for a lab-in-the-field experiment: 2,280 school-age students from low- to middle-income families, spanning grades 1 to 12 in the Spanish equivalent of the US education system. Participants came from four schools located in different cities across the southwestern region of Spain. All schools belong to Salesianos, a private network of Catholic schools funded by the Spanish government.[6] The experiment was integrated into the school's academic curriculum, with our Institutional Review Board (IRB) approving an opt-out procedure. No student opted out, so selection through opt-out is not a concern; the only students who did not participate were those absent on the day of the experiment (5.7%). Treatment assignment was implemented at the class or half-class level (rather than at the individual level), while pairings within each session were anonymous and randomly determined. This approach reflects the realities of school-based implementation: students are scheduled and tested one class at a time, making class-level assignment the feasible and standard unit of randomization. Moreover, individual-level assignment is difficult to communicate to young children and could introduce confusion or demand effects. Table 1 summarizes the distribution of participants by grade, type of competition, and token allocation method.[7]

[6] The Salesianos educational network has 59 schools in Spain serving almost 40,000 students (see `https://www.salesianos.es/escuelas`).

[7] Note that 11 participants failed to report demographics and 14 subjects failed to complete the non-cognitive tasks. Also, the initial-allocation regressions use 2,267 observations, whereas the final-allocation regressions use 2,255. The difference reflects 12 additional records in *addition* with an impossible final

Grade	1	2	3	4	5	6	7	8	9	10	11	12	Total
addition / effort	28	30	32	31	20	24	54	27	23	50	23	0	342
addition / talent	22	39	25	23	52	47	30	29	51	25	31	24	398
addition / luck	36	22	22	24	24	50	26	80	54	46	29	45	458
redistribution / effort	19	28	23	25	45	22	54	52	31	28	23	34	384
redistribution / talent	32	8	22	21	49	43	27	26	31	25	29	20	333
redistribution / luck	15	22	33	33	25	26	62	29	28	43	23	26	365
Total	152	149	157	157	215	212	253	243	218	217	158	149	2280

Table 1: Number of participants by grade, treatment, and condition

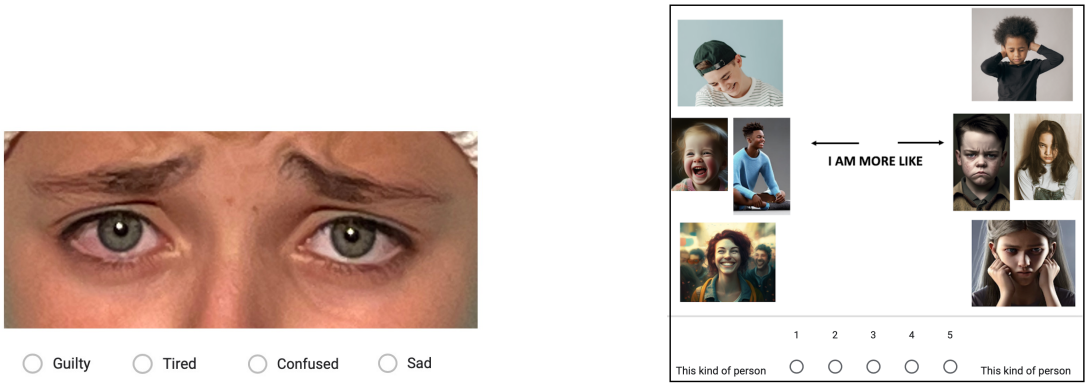
Payments. We identified a payment method that was uniform across all ages and highly valued by participants. All four schools in this sample have integrated cafeterias where students can purchase drinks, food and snacks at discounted prices. The cafeteria is very popular and widely accessed during recess and lunch, with items such as sandwiches and soft drinks highly and equally valued by students of all ages. To incentivize participants, we created our own currency, which could be spent on any item in the cafeteria. The currency had no expiration date, although most participants spent their endowments within a week. The exchange rate was 1 token = 1€.

Developmental Read the Mind in the Eyes Task (DeRMET). Next, we introduced participants to a novel, incentivized affective theory-of-mind task to examine the link between social preferences and the ability to empathize with others. This task builds on the original "Read the Mind in the Eyes Task" (RMET), a psychological assessment developed by Baron-Cohen et al. (1997) to measure individuals' ability to attribute complex mental and emotional states to others. In that task, participants view images of a person's eyes and select the adjective that best describes the emotion being conveyed. The RMET was originally designed for adults, with emotions that reflect subtle nuances and are described using sophisticated language. The standard adaptation of the RMET to children, which consists in presenting a subset of the original images (Baron-Cohen et al., 2001), has been shown to not perform very well in young populations (van der Meulen et al., 2017). We therefore created an alternative RMET task, called DeRMET (Developmental RMET), specifically for children and adolescents. The task was tested, registered with the Open Science Framework (OSF), pre-validated with 43 teenagers and adults, and validated with the 2,280 observations collected in this experiment (see Appendix B7 for details). It is intended as a tool for studying the development of emotional intelligence and its association with decision-making.

Our adaptation features 16 images capturing the expressive eyes of a 12-year-old pre-

allocation of zero. We treat these as recording errors and exclude them from the final-allocation analysis rather than recoding them.

teen who conveys 16 fundamental emotions, all described in simple language. For each image, we provide four possible answers, one of which most accurately represents the intended emotion. This modified version of the task offers three advantages for our population. First, sourcing all images from a single individual ensured greater consistency. Second, using a child as the model made the images more relatable for children and adolescents. Finally, focusing on simpler emotions and using straightforward language made the task appropriate for all ages, including the youngest participants.[8] Figure 3a shows one example image along with the four answer choices presented to participants (65% selected the correct answer, "sad").[9] Detailed instructions, including the full set of images, possible answers, correct responses, and the percentage of participants who answered each one correctly, are provided in Appendix A2.



(a) Developmental RMET **(b)** Pictorial Big Five (neuroticism trait)

Figure 3: Examples of DeRMET and P-Big 5

Pictorial Big Five (P-Big 5). Next, we administered a novel version of the Big Five questionnaire, to study the joint evolution of personality and social preferences. A typical objective in psychology is to categorize individuals across the "Big 5" personality traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism (Fiske, 1949; McCrae and Costa, 1987; Goldberg, 1993). Studies have employed a wide array of verbal questionnaires to examine those traits and how they relate to behavior and life outcomes. These verbal questionnaires can pose significant challenges for young participants, particularly when subtle emotional states are involved. In studies with children, assessments often rely on variants completed by parents or caregivers. However, for the

[8] Other authors have developed variations of the RMET task, or emotional intelligence tasks based on facial expressions (see e.g., Mattek et al. (2021); Pahnke et al. (2020)). To our knowledge, none are designed to study the developmental trajectory of affective theory-of-mind.

[9] An answer is labeled "correct" if it is *both* the emotion the preteen intended to convey in the image and the one chosen most often by participants. The two criteria coincide for all 16 images.

developmental trajectory of personality, it is crucial to employ a single assessment tool to ensure meaningful comparisons across various age groups.

To address these challenges and ensure consistent testing across individuals aged 5 to 18, we again developed a novel, child-friendly version of the Big 5 questionnaire, which we call the Pictorial Big 5 (P-Big 5). In this version, each trait is represented by a set of images illustrating opposite ends of the trait spectrum. Participants are shown a 5-point Likert scale and asked to select the option that best describes how they perceive themselves. The task was again tested, registered with OSF and pre-validated with 40 teenagers and adults who completed our P-Big 5 and a 25-question variant of the Five-Factor Model created for adolescents (Rogers and Glendon, 2018). We then validated the task with the 2,280 observations collected in this experiment (see Appendix B8 for details).

Our P-Big 5 questionnaire offers two advantages. Firstly, it does not rely on verbal comprehension, making it suitable for very young children and individuals from diverse linguistic backgrounds. Secondly, each question presents a natural choice between two polar ends, simplifying the decision-making process.[10] The primary limitation is that we offer just one question for each personality trait. This allows no space for cross-validation and focuses only on a specific facet of each dimension. Figure 3b presents the images employed for evaluating the "neuroticism" trait, and Appendix A3 provides the complete personality test, along with instructions and a comprehension example.

Questionnaires. After the final allocation decision, we asked participants in the effort and talent treatments to rate the difficulty of the competition task: 'Hard', 'Medium', 'Easy', or 'I don't know'. We also asked them how they thought their performance compared to their competitor's: 'Better', 'Equal', 'Worse', or 'I don't know'. These questions were not incentivized. Participants in the luck treatment were not asked these questions. We chose this timing intentionally to mitigate any potential demand effect that might prompt participants to alter their allocation in response to the task questions. At the end of the experiment we collected information regarding 'gender', 'date of birth', 'number of siblings' and 'favorite subject at school'.

Remarks. While we considered having each participant play only the observer or the competition task, we deemed this approach impractical, as it would have substantially reduced the number of observations for our primary task of interest, the observer task. In addition, our pilot study revealed that informing children that they would experience both roles enabled them to better understand the perspectives of both the winner and the loser, improving their comprehension of what it feels like to be in each position. By conducting the observer task before the competition task and then allowing participants to revise the allocation, we can document how decisions change after task participation. This

[10]This contrasts with verbal questions that typically present only one extreme (e.g., "I Am interested in people") and ask the individual how much they agree with the statement. In our experience, children understand the question better if they can picture the two polar cases.

sequence does not include a repeat-choice-only control condition, so changes may reflect experience, additional reflection, or repeated measurement; we therefore avoid assigning them a uniquely causal interpretation.

3 Conceptual framework

We formalize the problem as follows. Two players compete in a task, with the winner receiving x_W^o and the loser receiving x_L^o, with $x_L^o < x_W^o$. An observer can intervene in two ways: by *redistributing* funds from the winner to the loser or by splitting an *additional* amount between the two. Ultimately, the final allocations are x_W for the winner and x_L for the loser, with the total amount being in both cases fixed at N, so that $x_W + x_L = N$.

The observer takes into account both social norms and winner's merit, as determined by the competition outcome. Their preference over final allocations is represented by the *normative utility* function $(x_W)^{m_z} \times (x_L)^{1-m_z}$, where $m_z \in [1/2, 1]$. The parameter m_z reflects the weight the observer assigns to the winner's final utility, which depends on the nature of the competition, $z \in \{luck, effort, talent\}$. This parameter ranges from equal weight for both players ($m_z = 1/2$) to assigning full weight to the winner ($m_z = 1$).

The observer also experiences a disutility when actively reducing a player's final payment, captured by the *procedural disutility* function: $-d \max\{x_W^o - x_W, 0\} - d \max\{x_L^o - x_L, 0\}$, with $d > 0$. Consequently, the observer's overall preference is expressed as:

$$U(x_W, x_L) = (x_W)^{m_z} \ (x_L)^{1-m_z} - d \max\{x_W^o - x_W, 0\} - d \max\{x_L^o - x_L, 0\}$$

The observer's maximization problem is:

$$(x_W^*, x_L^*) = \arg\max_{x_W, x_L} U(x_W, x_L) \quad \text{s.t.} \quad x_W + x_L = N.$$

As is well known, when the observer is solely concerned with the participants' final utilities ($d = 0$), the winner's share of the total allocation is linearly related to their weight: $x_W^* = m_z N$. In this case, the allocation method (*addition* or *redistribution*) has no impact on the outcome. Under our experimental parameters ($N = 8$; $(x_W^o, x_L^o) = (6, 2)$ in *redistribution* and $(x_W^o, x_L^o) = (3, 1)$ in *addition*; $x_W \in \{4, 5, 6\}$), there exist two cutoff values m_1 and m_2 that determine the optimal allocation to the winner when $d = 0$: $x_W^* = 4$ if $m_z \in [1/2, m_1)$, $x_W^* = 5$ if $m_z \in [m_1, m_2)$ and $x_W^* = 6$ if $m_z \in [m_2, 1]$, with $x_L^* = 8 - x_W^*$.[11] As the procedural disutility parameter, d, increases, the loser's final share (weakly) decreases under *redistribution* while it remains unchanged under *addition*.

In the model, experiencing the task allows the observer to recalibrate their normative utility by adjusting m_z. The observer has an initial expected perception $E[r]$ of the merit/difficulty/value of winning the competition task. After participating in it, they

[11] Formally, $m_1 = \ln(4/3)/\ln(5/3) \simeq 0.563$ and $m_2 = \ln(3/2)/\ln(9/5) \simeq 0.690$

reevaluate their perception to r, and update the weight using a simple reinforcement learning rule:

$$m'_z = m_z + \alpha(r - E[r])$$

where α is a parameter representing the adaptation rate. In this framework, if the observer comes to view the task as more (respectively, less) deserving of credit than they initially believed, they adjust the winner's weight upward (respectively, downward).

The model provides a disciplined way to organize the empirical analysis, and its key primitives offer interpretations of the experimental contrasts. Differences between allocation mechanisms (*redistribution* vs. *addition*) are consistent with variation in the procedural disutility term d. Differences across winner-determination procedures (*luck*, *effort*, *talent*) are informative about the average merit weight m_z. Within-subject revisions from initial to final choices can be represented as updating, $\alpha(r - E(r))$. Finally, cross-sectional variation in age and measured non-cognitive skills describes how behavior associated with these primitives varies with development and individual characteristics. We do not structurally estimate the primitives, and the empirical contrasts need not uniquely identify any one mechanism.

The framework is written for impartial third-party judgments without personal stakes. Paradigms with self-interested decision makers are not nested in it, since personal payoffs would add terms and parameters to the objective function and change incentives. Results from such paradigms, and from studies of adult impartial observers, remain useful benchmarks for our findings, and we return to them in the conclusion. They do not, however, speak directly to how the allocation mechanism and the source of inequality interact with age and non-cognitive skills, which is the question our design addresses.

4 Results

We examine how the winner-selection (*talent*, *effort*, or *luck*) and the allocation environment (*redistribution* or *addition*) shape decisions across student ages. This section has five parts. First, we report the main treatment effects. Second, we study age patterns. Third, we analyze the factors that lead students to revise their allocations. Fourth, we assess the role of non-cognitive skills. Last, we relate our results to our conceptual framework. Summary statistics are reported in Appendix B1. Because treatment was assigned at the session level (a class or half-class), we cluster standard errors by session (131 sessions) and include campus fixed effects (FE).

Although the overall sample is large for an in-person experiment with children, the six treatments crossed with twelve grades imply an average of 32 observations in a treatment-by-grade cell. Our main inferences therefore rely on pooled regressions rather than individual cell comparisons. Throughout, we report point estimates and session-clustered standard errors, including for statistically imprecise coefficients, and treat fine treatment-

by-age contrasts as suggestive. When we state that two conditions are not statistically distinguishable, we do not interpret this as evidence that their effects are identical.

4.1 Main treatment effects

Table 2 presents the results of OLS regressions, where the dependent variables are the initial and final allocations to the loser. As independent variables, we include *Age* in months and its quadratic term Age^2, to account for potential non-linearities. We include dummies for environment and condition, with *addition* and *luck* being the default categories, respectively. Finally, we also control for gender ($Male = 1$), whether the participant has at least one sibling ($Siblings = 1$), and whether the participant's favorite school subject falls within STEM fields ($STEM = 1$), as a proxy for self-reported analytical inclination.

	Initial	Initial	Initial	Initial	Final	Final	Final	Final
Age	0.001	0.001	0.003	0.003	0.001	0.001	0.012**	0.012**
	(0.001)	(0.001)	(0.006)	(0.006)	(0.001)	(0.001)	(0.006)	(0.006)
Age^2			−0.000	−0.000			−0.000**	−0.000**
			(0.000)	(0.000)			(0.000)	(0.000)
talent	−0.059	−0.160*	−0.060	−0.162*	−0.165***	−0.132	−0.170***	−0.141
	(0.069)	(0.092)	(0.070)	(0.092)	(0.062)	(0.091)	(0.060)	(0.087)
effort	−0.081	−0.091	−0.082	−0.092	0.009	0.098	0.002	0.089
	(0.069)	(0.092)	(0.069)	(0.091)	(0.074)	(0.113)	(0.073)	(0.110)
redist.	−0.160***	−0.241**	−0.160***	−0.243**	−0.113**	−0.032	−0.114**	−0.038
	(0.058)	(0.097)	(0.058)	(0.097)	(0.054)	(0.094)	(0.054)	(0.092)
Male	−0.139***	−0.137***	−0.140***	−0.138***	−0.152***	−0.154***	−0.156***	−0.158***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
Siblings	−0.035	−0.037	−0.036	−0.038	−0.029	−0.030	−0.035	−0.035
	(0.046)	(0.046)	(0.046)	(0.046)	(0.048)	(0.047)	(0.046)	(0.045)
STEM	0.006	0.005	0.007	0.006	0.014	0.011	0.022	0.018
	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)	(0.033)	(0.034)
redist.×*talent*		0.224		0.226		−0.072		−0.063
		(0.138)		(0.138)		(0.117)		(0.117)
redist.×*effort*		0.032		0.033		−0.181		−0.176
		(0.136)		(0.136)		(0.148)		(0.147)
Num. obs.	2267	2267	2267	2267	2267	2267	2267	2267
R^2	0.024	0.028	0.024	0.028	0.030	0.032	0.035	0.037
Adj. R^2	0.019	0.023	0.019	0.022	0.025	0.027	0.030	0.031

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Campus FE (4 campuses). SE clustered at the session level (131 sessions).

Table 2: Main treatment effects in the *initial* and *final* allocation to the loser.

Across both decisions, allocations to the loser are lower in the *redistribution* environment than in the *addition* environment. This pattern complements evidence from adult dictator games showing a similar effect in taking vs. giving (List, 2007). Unlike List's manipulation, our comparison holds the feasible final allocations fixed and assigns the decision to an *impartial observer*. Because the additional tokens are initially unassigned, observers may nevertheless perceive them as an independent windfall rather than as part

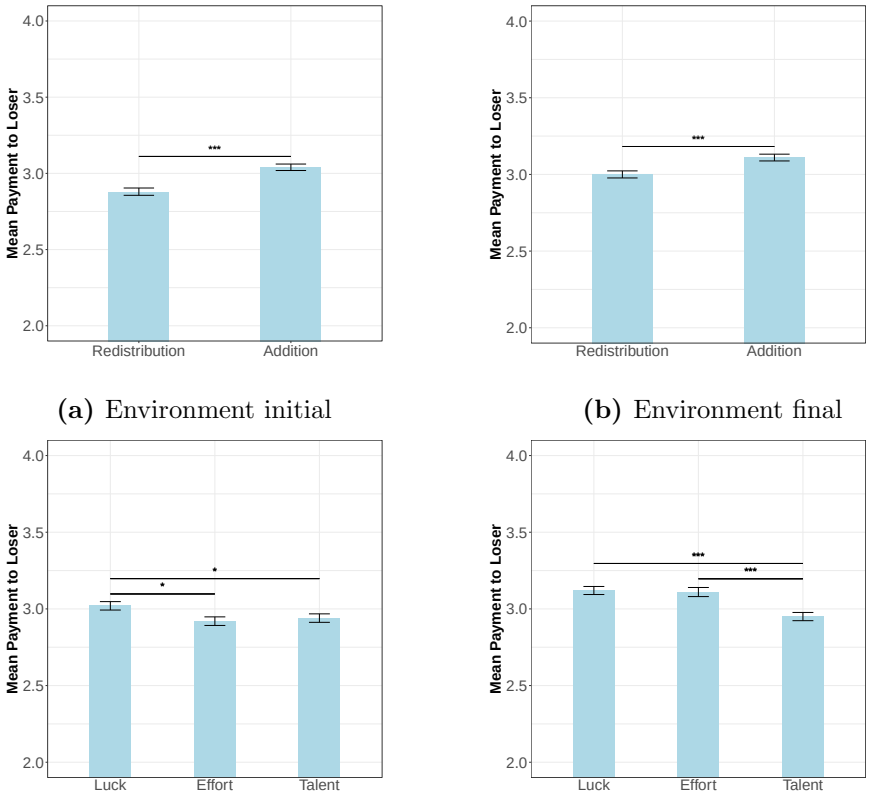
of the competition prize; we therefore interpret the comparison as the effect of the allocation procedures as implemented, not as uniquely identifying loss aversion or any other single mechanism. In the *initial* allocation, redistribution reduces transfers to the loser by 0.160 points (SE = 0.058) in the baseline specification and by 0.241 points (SE = 0.097) in the specification with environment-by-condition interactions. The interaction terms are imprecisely estimated, so we do not detect differential redistribution effects across *luck*, *talent*, and *effort*. In the *final* allocation, the corresponding baseline estimate is −0.113 (SE = 0.054); with interactions, the reference-condition estimate is −0.032 (SE = 0.094). Overall, the clearest evidence of an environment effect appears in the *initial* allocation. The result is also illustrated in Figures 4a and 4b.

We expected that allocations to the loser would be lowest under *talent*, intermediate under *effort*, and highest under *luck*. The estimates provide partial support for this ranking. Relative to *luck*, the pooled baseline coefficients for *talent* are −0.059 (SE = 0.069) in the initial allocation and −0.165 (SE = 0.062) in the final allocation. The corresponding coefficients for *effort* are −0.081 (SE = 0.069) and 0.009 (SE = 0.074). Thus, the talent-luck contrast is clearest in the pooled final-allocation specification, while the effort-luck estimates are imprecise. A direct comparison of the two conditions points the same way: the talent and effort coefficients are indistinguishable in the initial allocation (a difference of 0.022; Wald test, $p = 0.751$), whereas in the final allocation the talent coefficient lies 0.174 below the effort coefficient (Wald test, $p = 0.011$). The environment-by-condition interaction terms are also imprecisely estimated in both decisions, so we do not find strong evidence that either contrast differs between *addition* and *redistribution*. These estimates suggest that participants are more accepting of inequality when success is framed as *talent* (knowledge/ability). The slider task may have been viewed as monotonous or noisy and therefore less diagnostic of effort, whereas the trivia contest resembled a familiar test requiring knowledge. Raw effects are illustrated in Figures 4c and 4d.

Consistent with the literature on inequality acceptance in adults (Almås et al., 2020), we observe gender effects, with males exhibiting a stronger preference for meritocracy than females ($p < 0.01$ in all specifications of Table 2). This result contrasts with the findings of Almås et al. (2010), who studied a related but different game with children and adolescents and reported no gender differences in inequality acceptance. We report a further analysis of gender effects by interacting gender with environment and condition variables in Appendix B2. Most interaction terms are small or imprecisely estimated; the exceptions are discussed there.

Additionally, our expectation that children with siblings would allocate more to the losing party due to their familiarity with adult-induced redistribution was not supported by the data. Finally, we find no evidence that analytical orientation influences redistribution decisions.[12] We also note a concave age pattern in the final allocation, which we examine

[12]We interacted these variables with treatments and found no effect. These regressions are omitted for



(a) Environment initial **(b)** Environment final

(c) Condition initial **(d)** Condition final

Figure 4: Treatment effects

in the next section.

Result 1 *Main treatment estimates. Allocations to the winner are higher:*

- ***a)*** *In the redistribution environment.*
- ***b)*** *In the talent condition relative to luck, most clearly in the final allocation; the effort-luck contrast is imprecisely estimated.*
- ***c)*** *By male participants.*

conciseness but can be shared upon request.

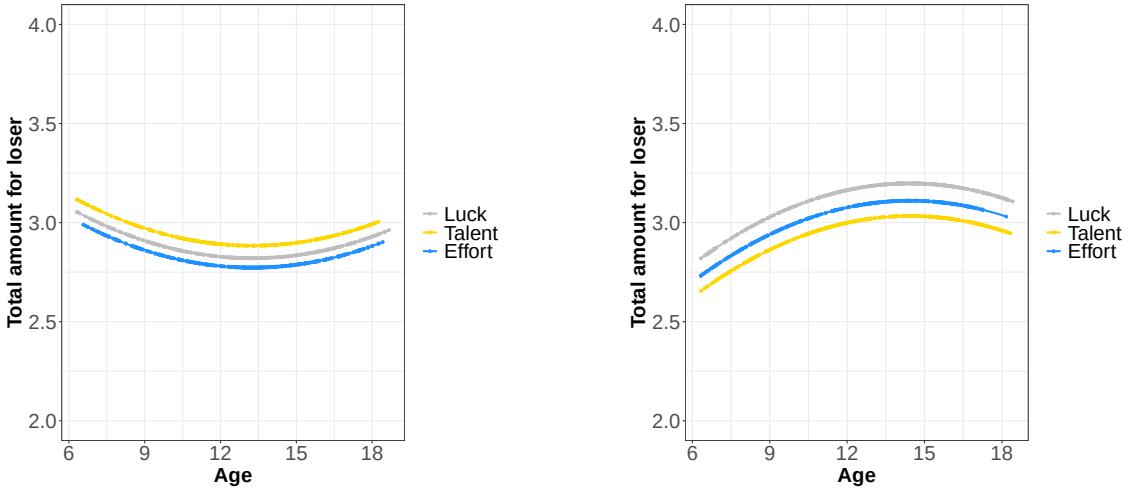
4.2 Age effects

4.2.1 Initial allocations

We now examine age effects in more detail. In Table 3, we report OLS regressions where the dependent variable is the *initial allocation to the loser*, and each treatment is separated into its condition and environment components, including interaction terms with age variables. Figures 5a and 5b display the quadratic best-fit curves for the initial allocation, plotted as a function of age and winner selection method, separately for the *redistribution* (left) and *addition* (right) environments.

In the environment-specific regressions (columns 1-4), the estimated developmental patterns differ across environments. In *redistribution*, allocations to the loser are essentially flat with age in both the linear and quadratic specifications. In *addition*, by contrast, allocations to the loser increase with age in the linear specification. In some specifications that allow for curvature, this age pattern is better described as concave, though the strength of this curvature depends on whether campus fixed effects are included and on the clustering assumptions (see Appendix B3). We therefore emphasize the difference between the mechanism-specific profiles rather than a precise functional form for the initial allocation.

The fitted curves in Figures 5a and 5b also suggest that the gap between *addition* and *redistribution* opens up from roughly grade 4 onward, with the youngest participants allocating slightly more to the loser under *redistribution*.



(a) Redistribution **(b)** Addition

Figure 5: *Initial* allocation to the loser across ages

	Redistribution	Redistribution	Addition	Addition	All	All
Age	−0.001	−0.009	0.002**	0.012	0.002**	0.013*
	(0.001)	(0.008)	(0.001)	(0.007)	(0.001)	(0.007)
Age^2		0.000		−0.000		−0.000
		(0.000)		(0.000)		(0.000)
talent	0.047	0.048	−0.130	−0.139	−0.054	−0.058
	(0.098)	(0.096)	(0.089)	(0.088)	(0.069)	(0.067)
effort	−0.063	−0.060	−0.061	−0.069	−0.072	−0.074
	(0.097)	(0.098)	(0.091)	(0.089)	(0.068)	(0.066)
Male	−0.110**	−0.106**	−0.158***	−0.159***	−0.141***	−0.138***
	(0.045)	(0.045)	(0.052)	(0.052)	(0.035)	(0.034)
Siblings	−0.053	−0.049	−0.026	−0.031	−0.033	−0.035
	(0.069)	(0.069)	(0.061)	(0.060)	(0.045)	(0.045)
STEM	0.017	0.012	−0.003	0.003	0.005	0.007
	(0.046)	(0.046)	(0.044)	(0.043)	(0.033)	(0.032)
redist.					0.276	1.751**
					(0.199)	(0.685)
Age×*redist.*					−0.003**	−0.024**
					(0.001)	(0.010)
Age^2×*redist.*						0.000**
						(0.000)
Num. obs.	1078	1078	1189	1189	2267	2267
Num. groups FE	4	4	4	4	4	4
Num. clusters SE	62	62	69	69	131	131
R^2	0.013	0.016	0.041	0.045	0.029	0.034
Adj. R^2	0.005	0.007	0.034	0.037	0.024	0.029

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 3: *Initial* allocation to the loser by treatments and age

To test whether developmental profiles differ across environments, columns 5-6 in Table 3 pool the data and interact *Age* (and, in column 6, Age^2) with the *redistribution* indicator. In these pooled models, the coefficient on *Age* (and Age^2) captures the developmental profile in the *addition* environment (the reference environment), while the interaction terms indicate how the *redistribution* profile departs from it. The negative and statistically significant *Age*×*redistribution* interaction indicates that the age-related increase estimated under *addition* is weaker under *redistribution*. In the quadratic specification, the interaction estimates are also consistent with different curvature across environments. Because *Age* is measured in months, the main effect of *redistribution* in the pooled specifications reflects differences evaluated at the smallest age and is not directly interpretable; this is why we focus on the implied age profiles and interaction terms rather than on the redistribution intercept.[13]

[13] Appendix B3 shows that the qualitative distinction between a comparatively flat profile in *redistribution*

Result 2 *For the initial choices:*

- ***a)*** *The estimated allocation to losers is comparatively flat with age under redistribution and increasing under addition. Evidence on concavity in the initial allocation depends on the econometric specification.*
- ***b)*** *The fitted profiles suggest that the tendency to allocate more resources to losers under addition than under redistribution appears from roughly grade 4 onward.*
- ***c)*** *Male participants allocate fewer resources to losers.*

4.2.2 Final allocations

Table 4 and Figure 6 report the analogue of Table 3 and Figure 5 for the *final* decision. As before, Appendix B3 replicates the analysis under alternative specifications, and Appendix B4 replaces *Age* in months with *Age-group* indicators. Recall that between the initial and final choices, observers experienced the tournament, which allowed them to assess the task's difficulty and to form an expectation about their own performance relative to their competitor. However, they were not informed of their actual outcome, and the design did not allow for any form of reciprocity. And yet, one might expect that experiencing the task could influence behavior, for example, by increasing empathy toward the loser if the participant perceived it as difficult or believed they underperformed. In practice, approximately two-thirds of participants did not revise their initial allocation.[14]

In *redistribution*, final allocations display the same developmental pattern as initial choices: neither *Age* nor Age^2 significantly predict allocations to the loser in the environment-specific regressions. In contrast with the initial decision, however, the *talent* condition is now associated with a sizable reduction in the loser's final allocation relative to *luck*, whereas *effort* remains statistically indistinguishable from *luck*. The pooled specifications with age-by-environment interactions give the same picture.

In *addition*, the final allocation increases with age in the linear specification and is concave in the quadratic specification. Regarding winner-selection conditions, the point estimates for *talent* and *effort* (relative to *luck*) are smaller and less precisely estimated within addition than within redistribution, while the pooled regressions indicate that, on average, *talent* is associated with lower allocations to the loser.

The last two columns pool both environments and test whether developmental profiles differ across *addition* and *redistribution* by interacting the age variables with the redistribution indicator. In these regressions, the coefficients on *Age* and Age^2 describe the age profile in the *addition* environment (the reference), while the interaction terms

and an increasing, sometimes concave, profile in *addition* appears across several alternative specifications. The exact curvature and its precision are specification-dependent. Appendix B4 reports estimates using age-group indicators.

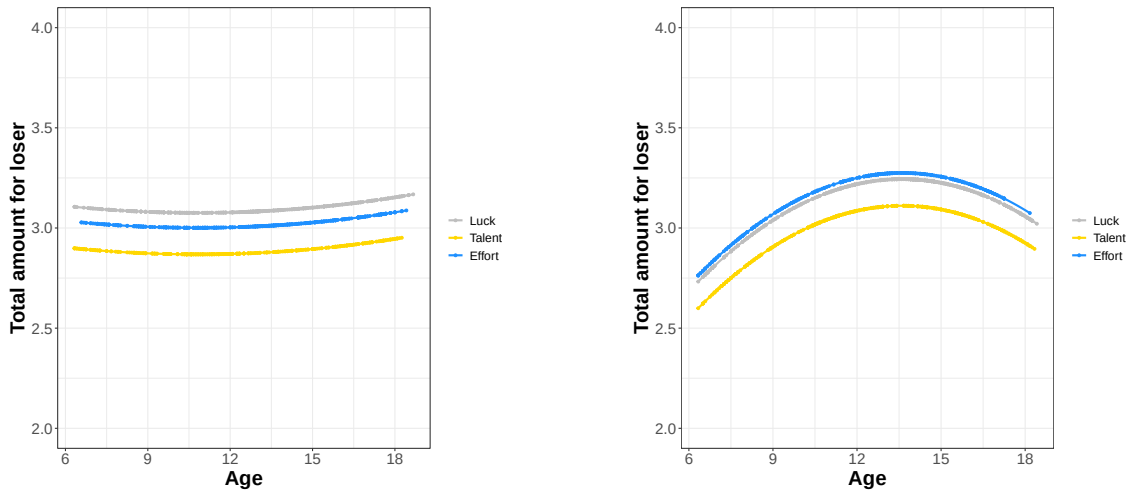
[14] In *addition*, 79% (luck), 69% (effort) and 67% (talent) of participants did not revise their initial allocation. In *redistribution*, 65% (luck), 61% (effort) and 65% (talent) did not revise it.

	Redistribution	Redistribution	Addition	Addition	All	All
Age	0.001	0.003	0.003^{**}	0.021^{***}	0.002^{**}	0.024^{***}
	(0.001)	(0.007)	(0.001)	(0.008)	(0.001)	(0.008)
Age^2		−0.000		-0.000^{**}		-0.000^{***}
		(0.000)		(0.000)		(0.000)
talent	-0.215^{***}	-0.215^{***}	−0.089	−0.105	-0.155^{**}	-0.164^{***}
	(0.075)	(0.076)	(0.090)	(0.082)	(0.062)	(0.058)
effort	−0.084	−0.085	0.071	0.054	−0.011	−0.020
	(0.091)	(0.091)	(0.096)	(0.089)	(0.070)	(0.066)
Male	-0.106^{*}	-0.107^{*}	-0.201^{***}	-0.203^{***}	-0.157^{***}	-0.157^{***}
	(0.056)	(0.056)	(0.043)	(0.043)	(0.035)	(0.035)
Siblings	−0.020	−0.021	−0.059	−0.068	−0.036	−0.043
	(0.058)	(0.058)	(0.071)	(0.066)	(0.047)	(0.044)
STEM	0.022	0.023	0.003	0.016	0.010	0.018
	(0.051)	(0.051)	(0.044)	(0.044)	(0.034)	(0.033)
redist.					0.178	1.680^{**}
					(0.221)	(0.761)
Age×*redist.*					−0.002	-0.024^{**}
					(0.001)	(0.010)
Age^2×*redist.*						0.000^{**}
						(0.000)
Num. obs.	1078	1078	1177	1177	2255	2255
Num. groups FE	4	4	4	4	4	4
Num. clusters SE	62	62	69	69	131	131
R^2	0.028	0.028	0.049	0.064	0.032	0.042
Adj. R^2	0.019	0.019	0.042	0.056	0.027	0.037

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 4: *Final* allocation to the loser by treatments and age

capture how the *redistribution* profile departs from it. The negative and statistically significant *Age*×redistribution interaction, together with the offsetting curvature term in the quadratic specification, indicates that the age-related increase and concavity estimated under addition are attenuated under redistribution, consistent with a near-flat redistribution profile. Again, because *Age* is measured in months, the standalone redistribution coefficient corresponds to the difference evaluated at the smallest age and is not directly interpretable; we focus on the implied profiles and interaction terms. Appendix B3 reports qualitatively similar estimates when campus fixed effects are omitted. With two-way clustering at the session and campus levels, standard errors are larger given the small number of campuses and the interaction terms lose statistical significance, while the point estimates are unchanged. As before, male participants display more meritocratic preferences. These patterns are broadly similar when we run separate regressions for each winner selection method (see Appendices B5 and B6 for details).



(a) Redistribution **(b)** Addition

Figure 6: *Final* allocation to the loser across ages

Result 3 *For choices made after experiencing the competition task:*

- ***a)*** *Estimated final transfers to losers are comparatively flat with age under redistribution, whereas the quadratic specification under addition peaks around age thirteen.*
- ***b)*** *Point estimates suggest a stronger winner-favoring pattern in the talent condition among participants in grade 4 and above , but these finer age-by-condition comparisons are less precise.*
- ***c)*** *Male participants remain more likely to favor the winner than female participants.*

4.3 What makes subjects revise their initial decision?

Final choices are made after participants have experienced the competition task and have been invited to reconsider their initial allocation. We next document correlates of the shifts in choices. These estimates do not isolate a causal effect of competitive experience from reflection, reconsideration, or repeated measurement.

During the experiment, we included two sets of non-incentivized questions regarding the participants' perceptions about the level of difficulty of the competition task and their performance relative to others. These questions were not included in the *luck* condition, as the winner was randomly determined. We expected to observe self-serving biases among participants: those who perceived their performance as higher than others and found the task easy would likely lean towards more meritocratic allocations (and vice versa).[15]

[15] One must acknowledge the exogenous limits to this tendency, as individuals with corner initial choices, (4,4) and (6,2), cannot shift towards more and less redistribution, respectively.

For our analysis, we consider each allocation method and each initial choice separately: meritocratic (6,2), equitable (5,3) and compensatory (4,4). For initial meritocratic (Merit.) and compensatory (Comp.) choices, we perform Poisson regressions, where we interpret the absolute value of the change in the allocation to the winner between the initial and final choice as counts $(0, 1, 2)$. For initial equitable choices (Equit.), we perform a multinomial regression of the likelihood to decrease (-1) or increase $(+1)$ the final allocation of the winner. Thus, a positive coefficient in Merit. and Equit.(-1) implies an increase in the allocation to the loser whereas a positive coefficient in Comp. and Equit.$(+1)$ implies an increase in the allocation to the winner. We include the same controls as before, exclude the *luck* condition, and pool the *effort* and *talent* data with a *talent* dummy.

Regarding our variables of interest, we use "equal / I don't know" as the reference category for self-reported perceived performance and introduce dummies for *better* and *worse*. Similarly, for self-reported perceived difficulty, we set "same / I don't know" as the reference and include dummies for *hard* and *easy*. Table 5 presents the results.

Perceived own performance is the strongest predictor of revisions. Participants who believed they had *performed better* than their competitor revised toward the winner: they moved away from an initial meritocratic choice less often, moved from equitable to compensatory less often, moved from equitable to meritocratic more often and, under *redistribution*, moved away from an initial compensatory choice more often (the corresponding coefficient under *addition* has the opposite sign and is imprecise). Participants who believed they had *performed worse* revised in the opposite direction; these estimates are precise under *addition* and imprecise under *redistribution*.[16] Perceived difficulty has weaker and less consistent associations with revisions. An *easy* task is associated with fewer moves from equitable to compensatory under *redistribution* and with more upward revisions from compensatory choices under *addition*, both in the self-serving direction, but the coefficients on *hard* have mixed signs across columns. This pattern is consistent with an egocentric interpretation and with some existing findings in the literature. For example, Neuber (2021) found that individuals engaged in distributing rewards to others often adhere to norms that favor their own interests, even in situations where their choices exclusively impact others. Since perceived performance and difficulty are self-reported, however, the estimates should be read as associations rather than as causal effects.

Result 4 *Revisions are consistent with a self-serving pattern: individuals who perceived their performance as above their competitor's, and to a lesser extent those who considered the task easy, revised allocations in favor of the winner.*

[16] No participant in *addition* who reported performing worse moved from the equitable to the meritocratic allocation, so the corresponding coefficient in Table 5 is not identified and its standard error is not meaningful.

	redistribution				*addition*			
	Merit.	Equit. (−1)	Equit. (+1)	Comp.	Merit.	Equit. (−1)	Equit. (+1)	Comp.
Age	−0.013**	−0.854*	−0.839*	0.017	0.014	−0.218	−0.774***	−0.038**
	(0.005)	(0.499)	(0.448)	(0.063)	(0.090)	(0.530)	(0.270)	(0.016)
Age^2	0.0004**	−0.030	0.030*	−0.0001	−0.0001	0.007	0.030***	0.0001***
	(0.0002)	(0.019)	(0.018)	(0.0001)	(0.0001)	(0.022)	(0.009)	(0.0001)
Male	−0.314**	0.610	0.464	0.139	−0.428**	−0.049	0.495	0.510***
	(0.189)	(0.376)	(0.400)	(0.376)	(0.209)	(0.062)	(0.321)	(0.159)
Siblings	0.068	1.167	−0.187	−0.141	0.027	−0.388	−0.312	0.571**
	(0.085)	(0.793)	(0.545)	(0.183)	(0.166)	(0.403)	(0.325)	(0.275)
STEM	0.183	−0.628*	−0.662	0.057	0.134	−0.007	0.689***	−0.297*
	(0.214)	(0.381)	(0.413)	(0.225)	(0.212)	(0.154)	(0.099)	(0.164)
Talent	−0.139	−1.242***	−0.455	0.465	−0.526*	0.066	0.112	0.039
	(0.227)	(0.417)	(0.404)	(0.323)	(0.269)	(0.234)	(0.393)	(0.422)
Better	−0.422**	−0.921*	0.757*	0.543**	−0.856***	−0.209***	0.720***	−0.324
	(0.198)	(0.552)	(0.401)	(0.277)	(0.241)	(0.070)	(0.118)	(0.308)
Worse	0.322	0.424	−0.639	−1.123	0.331**	1.457***	−13.562***	−0.817*
	(0.212)	(0.449)	(0.811)	(0.839)	(0.149)	(0.183)	(0.000)	(0.431)
Hard	0.379***	−1.380	−0.341	0.586*	−0.110	0.918**	−0.977	−0.390
	(0.064)	(0.859)	(0.848)	(0.329)	(0.424)	(0.743)	(0.676)	(0.772)
Easy	0.037	−0.839*	0.227	0.187	−0.222	0.031	0.384	0.526***
	(0.242)	(0.456)	(0.400)	(0.330)	(0.158)	(0.340)	(0.400)	(0.104)
Constant	0.328	4.222	4.169	−2.312	−0.672	0.379	2.157	0.382
	(0.392)	(3.302)	(2.920)	(1.521)	(1.591)	(3.231)	(1.894)	(1.349)
Obs.	266	272	272	166	191	353	353	177
AIC	526	443	443	321	343	573	573	257

*p<0.1; **p<0.05; ***p<0.01

Table 5: Poisson and multinomial regressions of change in allocation to winner as a function of initial choice (meritocratic, equitable, compensatory) by treatment.

4.4 Non-cognitive abilities and decision-making

Recent research has found that cognitive abilities shape our time and risk preferences (Dohmen et al., 2010), although the exact mechanism is still not fully understood (Amador-Hidalgo et al., 2021). In this study we take a different route and hypothesize a link between non-cognitive abilities –empathy and personality– and social preferences. Empathy and personality are measured with the DeRMET and P-Big 5 tasks described in section 2. Performance by age is summarized in Figure 7 and discussed in Appendices B7 (DeRMET) and B8 (P-Big 5), which also report validation and comparative analyses.

In line with the existing literature, we find sustained and statistically significant changes in empathy and personality scores with age. Both tasks are also informative: the empathy test is neither trivial nor exceptionally difficult for any age, and the results of our pictorial personality test align well with the answers to the standard verbal

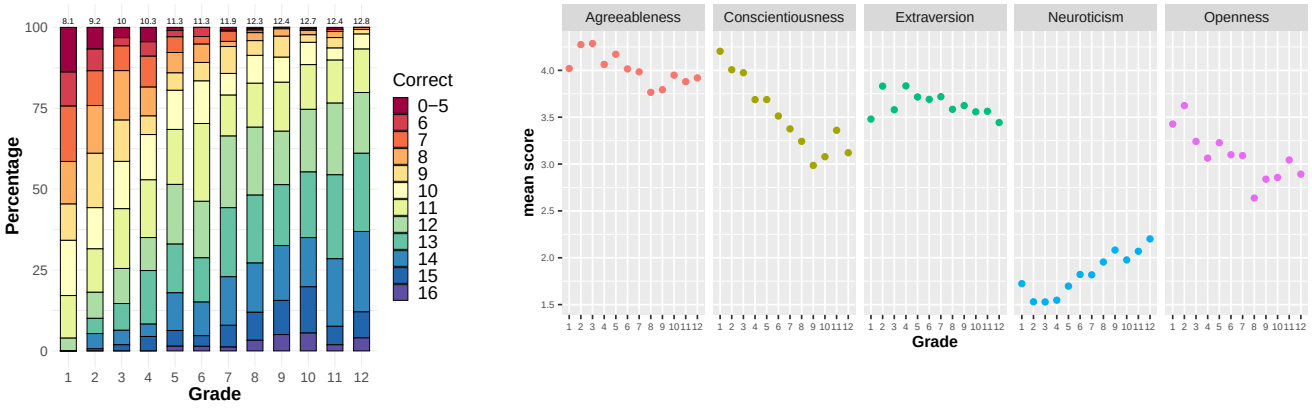


Figure 7: Left: Distribution of correct answers in DeRMET by grade with grade average on top. Right: Average score in the P-Big 5 by trait and grade

questionnaires (see Appendices B7 and B8).

To examine the interplay between the ability to interpret emotions, individual personality traits, and meritocratic decision-making, we estimate multinomial regressions using the "equitable" allocation (5,3) as the reference category. The dependent variables include "Merit." and "Comp.", that is, choosing instead (6,2) and (4,4) respectively.

To enhance statistical power, we pool data from all variants. We include dummies for treatment (*redistribution*, with *addition* as the reference) and winner selection method (*luck* and *talent*, with *effort* as the reference), and the usual control variables (*Age*, Age^2, *Male*, *Siblings*, *STEM*). For each individual, we assess the number of correct answers in the DeRMET and their responses to the P-Big 5. To control for age-related trends and extract individual-specific traits, we first regress each score on age and use the residuals (*resid*) as explanatory variables. We carry out the regressions for both the initial and final choices.[17] The results are presented in Table 6.

Participants with higher DeRMET scores are less likely to choose the meritocratic allocation over the equitable one in the initial decision. This is consistent with a stronger ability to read others' emotions being associated with more generous treatment of the loser. The association is absent in the final decision. One reading is that perceptions of own performance, which are strongly associated with revisions (Section 4.3), displace the initial emotional consideration; our design does not allow us to test this directly.

Personality traits are also associated with decisions. *Conscientiousness* is associated with a greater likelihood of meritocratic choices in the final decision, and *agreeableness* with a lower likelihood in both the initial and final decisions. *Openness* is associated with compensatory choices, mainly in the initial decision. These associations align with

[17] We also ran the analysis using standardized scores to reflect differences in scores relative to peers in the same grade. The results are very similar in both magnitude and significance of effects.

	Initial choice		Final choice	
	Merit.	Comp.	Merit.	Comp.
Age	-0.321^{***}	-0.308^{***}	-0.846^{***}	-0.415^{***}
	(0.091)	(0.077)	(0.215)	(0.108)
Age^2	0.011^{***}	0.012^{***}	0.031^{***}	0.016^{***}
	(0.004)	(0.003)	(0.008)	(0.004)
Male	0.472^{***}	0.062	0.680^{***}	0.019
	(0.068)	(0.134)	(0.207)	(0.115)
Siblings	−0.018	−0.137	0.009	−0.138
	(0.235)	(0.148)	(0.147)	(0.085)
STEM	0.0001	−0.021	-0.192^{*}	−0.122
	(0.074)	(0.054)	(0.104)	(0.100)
DeRMET (resid)	-0.059^{***}	0.003	−0.022	0.008
	(0.019)	(0.015)	(0.026)	(0.030)
Extraversion (resid)	−0.002	0.059^{*}	−0.063	0.030
	(0.025)	(0.033)	(0.077)	(0.031)
Agreeableness (resid)	-0.136^{**}	−0.082	-0.150^{***}	−0.064
	(0.064)	(0.054)	(0.052)	(0.098)
Conscientiousness (resid)	0.071	0.061	0.158^{**}	0.065
	(0.052)	(0.055)	(0.066)	(0.054)
Neuroticism (resid)	0.034	−0.008	0.055	−0.040
	(0.022)	(0.045)	(0.045)	(0.046)
Openness (resid)	0.028	0.109^{**}	−0.014	0.022
	(0.062)	(0.052)	(0.093)	(0.058)
redistribution	0.683^{***}	0.107	0.277^{***}	−0.096
	(0.070)	(0.178)	(0.081)	(0.177)
luck	0.110	0.460^{***}	−0.140	−0.027
	(0.245)	(0.151)	(0.125)	(0.109)
talent	−0.068	0.053	−0.044	-0.532^{***}
	(0.154)	(0.083)	(0.110)	(0.181)
Constant	2.078^{***}	0.864	4.943^{***}	2.472^{**}
	(0.369)	(0.670)	(0.950)	(1.004)
Obs.	2255	2255	2255	2255
AIC	4,730	4,730	4,664	4,664

*p<0.1; **p<0.05; ***p<0.01.

Table 6: Multinomial regressions of meritocratic and compensatory initial and final choices: the effect of emotional intelligence and personality

prior research linking conscientiousness to a preference for right-wing political parties and openness to a preference for left-wing political parties (Vecchione et al., 2011; Furnham and Fenton-O'Creevy, 2018; Ekstrom and Federico, 2019).

Allowing traits to jointly moderate the effects of the environment and the winner-selection condition, we find that most high-order interactions are imprecisely estimated. Statistically significant three-way interactions are few and do not point to a pervasive pattern, suggesting limited systematic heterogeneity beyond the main treatment contrasts

(see Appendix B9 for details).

Appendix B10 reports regressions of the change between initial and final choices with traits as additional explanatory variables. Performance in *DeRMET* is associated with upward revisions toward the winner in two of the eight specifications (initially compensatory choices under *redistribution* and initially equitable choices under *addition*), conditional on task perceptions. Personality traits are also related to revision patterns, conditional on perceived performance and difficulty. Higher *extraversion* is associated with persistence in winner-favoring allocations under *redistribution*. Higher *agreeableness*, higher *openness* and lower *conscientiousness* are associated with revisions towards the loser under *addition*.

Finally, we notice that the tendency for males to display a stronger preference for meritocratic allocations remains significant even after controlling for *DeRMET* performance and personality attributes.

Result 5

a) *Participants with higher emotional intelligence are less likely to initially exhibit meritocratic preferences.*

b) *As for personality, conscientiousness is associated with a greater likelihood of meritocratic final choices, and agreeableness with a lower likelihood of meritocratic choices.*

c) *Male participants are more likely to be meritocratic in both initial and final decisions.*

4.5 Crosswalk from empirical contrasts to model primitives

The conceptual framework is not introduced to "explain away" the data, but to organize possible interpretations of the experimental contrasts. The design varies two dimensions: token allocation environment and winner selection method. In the model, the first dimension is represented by a potential aversion to taking resources from parties (d), and the second by shifts in the merit weight (m_z). Within-subject revisions are represented by updating of the merit weight (m'_z). These interpretations require assumptions that the experiment does not separately test. Table 7 maps each empirical pattern to its corresponding model representation.

The results in sections 4.1, 4.2 and 4.3 offer the following interpretations. First, the allocation procedure matters: lower transfers to losers under *redistribution* are consistent with a positive procedural-disutility term in our model ($d > 0$), although, as noted in section 4.1, they are also consistent with observers treating the added tokens as a windfall. Second, point estimates suggest a higher merit weight under *talent* than under *luck*, especially for the final allocation. Third, the two estimated developmental patterns may reflect, within the model, a tension between fairness preferences and procedural disutility. Under *addition*, the procedural-disutility term is moot since $x_W > x_W^o$ and $x_L > x_L^o$ by design; under *redistribution*, it may offset developmental changes in fairness preferences. The developmental literature on fairness does not predict this divergence, and our design does

Empirical contrast / main finding	**Model primitive**	**Interpretation within the model**
Redistribution v. addition: loser receives less under redistribution (initial and final)	d	Consistent with procedural disutility $d > 0$ (taking compresses equalization).
Luck v. effort v. talent: final allocations are most meritocratic under talent; effort $\approx$ luck	m_z	Point estimates are consistent with a larger merit weight under talent ($m_T > m_E \approx m_L$).
Age gradients differ by mechanism: concavity in addition; flat profile in redistribution	m_z(age), d(age)	Mechanism-specific age profiles can be represented by age variation in m_z and d.
Within-subject revision: direction correlates with perceived performance	α	Revisions correlated with perceptions are consistent with updating: $m_z' = m_z + \alpha(r - E[r])$, with $\alpha > 0$.
Gender (Male): allocations are more meritocratic	m_z(gender)	Baseline heterogeneity in the merit weight: m_z(male) $> m_z$(female).
Non-cognitive traits: heterogeneity in allocations and/or moderation of treatment effects	m_z (and possibly d, α)	Shifts in baseline m_z and α; trait$\times$condition interactions are consistent with trait-dependent shifts in m_z across conditions; little evidence of trait-related variation in d.

Table 7: Crosswalk from empirical contrasts to model primitives

not identify the channel: separating it from alternatives would require varying the procedure within subject or eliciting procedural attitudes directly. Fourth, almost one-third of our subjects revise their initial allocation, and the correlation between these revisions and perceived performance is consistent with updating, among other explanations.

Regarding gender, the fact that males are systematically more meritocratic is most naturally read as a higher baseline merit weight for males, i.e. m_z(male) $> m_z$(female). This gap holds both in the initial and final choices, and remains when controlling for DeRMET and Big-5 traits. However, we do not detect differential sensitivity to the environment (the coefficient for the interaction *Male*$\times$*Redistribution* is small and insignificant). The significant *Effort*$\times$*Male* interaction in the final allocation is consistent with the effort condition shifting the implied merit weight differently for males and females, while we find no evidence that gender operates through d.

Last, non-cognitive traits are associated with heterogeneity in allocations and, to a limited extent, with the size of treatment effects. Empirically, higher DeRMET (emotional recognition/empathy) is associated with a lower propensity to choose the meritocratic split in the *initial* allocation, but this association largely vanishes in the *final* decision, which is consistent with, but not evidence of, experience-based updating between choices. In contrast, personality traits exhibit more persistent patterns: conscientiousness is associated with more meritocratic final choices (higher m), agreeableness with less meritocratic choices in both decisions (lower m), and openness with a greater tendency towards com-

pensatory allocations in the initial decision. The weak *trait*×*environment* interactions suggest that there is little evidence that d varies systematically with individual characteristics once we allow for rich interactions. Several *trait*×*condition* terms are statistically significant, which is consistent with trait-dependent shifts in the merit weight across conditions, although with this many interaction terms some significant coefficients are expected by chance.

Several traits are also associated with revision patterns. Extraversion, agreeableness, conscientiousness, and openness are the traits most consistently related to revisions across specifications, and higher DeRMET scores are associated with revisions toward the winner in some specifications. These results point to heterogeneity in revision behavior related to affective and personality traits.

5 Conclusion

In this study, impartial observers aged 5 to 18 in Spain decide the allocation of rewards between a winner and a loser. We examined three (between-subject) conditions for determining the winner, *effort*, *talent*, and *luck*; two (between-subject) methods for allocating rewards, *redistribution* and *addition*; and two (within-subject) timings for choice, before and after experiencing the task.

Participants in all age groups allocated less to the loser when equalization required actively reducing the winner's tokens. Under *addition,* the estimated age profile of final allocations rose through middle school and declined afterwards; under *redistribution*, it was comparatively flat. In pooled final-allocation specifications, observers allocated less to the loser when the winner was determined through *talent*. Revisions were associated with participants' perceptions of their own performance and, less consistently, of task difficulty; because all participants were invited to reconsider, revisions may also reflect reflection or repeated measurement. Individual characteristics and non-cognitive traits were also related to decisions. In particular, male participants exhibited more meritocratic allocations in both their initial and final choices.

Overall, the study jointly examines how responses to the origin of inequality and to redistribution mechanisms vary between childhood and adolescence. The estimates indicate an early sensitivity to both outcomes and procedures. A few additional remarks are in order.

The lower allocation to the loser under redistribution than under addition is consistent with an aversion to intervention, reminiscent of status quo bias (Kahneman et al., 1991). Our evidence suggests that procedural sensitivity is present in third-party judgments and among children and adolescents. One possible interpretation is observer-side loss aversion: taking from a winner may be coded as a loss-creating act. An alternative is that observers treat the added tokens as a windfall that is not part of the competition prize. The

experiment raises the possibility that institutionally equivalent allocations may be viewed differently depending on how they are implemented.

Our sample covers ages 5 to 18 and therefore speaks directly to childhood and adolescence, not to later stages of the life cycle. The closest adult benchmarks are the impartial-spectator studies of Cappelen et al. (2022) and Almås et al. (2020), in which adult observers equalize more often when inequality results from luck than from productivity. Among our oldest participants, the *talent* condition yields the lowest final allocation to the loser in both environments (Figure 6), which is the same direction; the designs differ too much to compare magnitudes. Evidence from self-interested dictator games indicates that age patterns can continue to change in adulthood and older age, although the direction depends on the paradigm. For example, Cobo-Reyes et al. (2019) report non-linear changes in egalitarian and altruistic types between ages 9 and 67, while the meta-analysis by Engel (2011) finds greater giving in older subject pools. These findings are useful benchmarks but are not directly comparable to impartial-observer choices. Whether the mechanism-specific profiles documented here persist, converge, or reverse in adulthood and old age remains an open question.

Between 20% and 40% of observers revised their choices after taking part in the competition task and being invited to reconsider their initial allocation. These revisions are systematically associated with participants' perceptions of their own performance, a pattern consistent with experiential learning (Kolb et al., 2014): task participation may provide information about difficulty and the requirements for success while also introducing subjective anchors. Similar patterns may arise in other decision-making paradigms where participants take on roles different from their own.

The talent-luck point estimates, clearest in the pooled final-allocation specifications, suggest a tendency to reward winners perceived as more deserving. The evidence is less precise for effort, and we therefore do not interpret the estimates as establishing a general ranking of talent, effort, and luck. The result nevertheless suggests that sensitivity to the perceived source of inequality may emerge early in life.

The importance of the source of inequality aligns with previous findings (Cappelen et al., 2010), and has been sometimes linked to political affiliations and individual characteristics (Alesina and Giuliano, 2011; Luttmer and Singhal, 2011; Almås et al., 2020). We are able to link individual heterogeneity to non-cognitive skills. Participants with a higher capacity for empathy make fewer merit-based initial choices, and those who score higher on agreeableness and openness and lower on conscientiousness lean towards greater redistribution.

Finally, males exhibit more merit-based choices than females. This observation could be linked to the documented tendency of women to be more generous than men in dictator games (Eckel and Grossman, 1998; Andreoni and Vesterlund, 2001; Brañas-Garza et al., 2018), except that these findings have not been consistently replicated in children (Fehr et al., 2013; Brocas et al., 2019; Brocas and Carrillo, 2020b). Although our study was not

specifically designed to investigate gender differences and we lack the appropriate tools to draw definitive conclusions, the robustness of the observed correlation suggests that males and females may approach the well-being of others differently. Interestingly, we also found that females demonstrated a greater ability to perceive and interpret others' emotions.

References

Alberto Alesina and Paola Giuliano. Preferences for redistribution. In *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier, 2011.

Antonio Alfonso, Pablo Brañas Garza, Isabelle Brocas, Juan D Carrillo, Juan B Gonzalez, and Maria Jose Vazquez. Reading minds to win: The power of cognitive and affective skills in children's strategic play. Work in progress, 2025.

Ingvild Almås, Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. Fairness and the development of inequality acceptance. *Science*, 328(5982):1176–1178, 2010.

Ingvild Almås, Alexander W Cappelen, and Bertil Tungodden. Cutthroat capitalism versus cuddly socialism: are americans more meritocratic and efficiency-seeking than scandinavians? *Journal of Political Economy*, 128(5):1753–1788, 2020.

Luis Amador-Hidalgo, Pablo Brañas Garza, Antonio M Espín, Teresa García-Muñoz, and Ana Hernández-Román. Cognitive abilities and risk-taking: Errors, not preferences. *European Economic Review*, 134:103694, 2021.

James Andreoni and Lise Vesterlund. Which is the fair sex? gender differences in altruism. *The Quarterly Journal of Economics*, 116(1):293–312, 2001.

James Andreoni, Deniz Aydin, Blake Barton, B Douglas Bernheim, and Jeffrey Naecker. When fair isn't fair: Understanding choice reversals involving social preferences. *Journal of Political Economy*, 128(5):1673–1711, 2020.

Silvia Angerer, E Glenn Dutcher, Daniela Glätzle-Rützler, Philipp Lergetporer, and Matthias Sutter. Gender differences in discrimination emerge early in life: Evidence from primary school children in a bilingual city. *Economics Letters*, 152:15–18, 2017.

Simon Baron-Cohen, Therese Jolliffe, Catherine Mortimore, and Mary Robertson. Another advanced test of theory of mind: Evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry*, 38(7):813–822, 1997.

Simon Baron-Cohen, Sally Wheelwright, Amanda Spong, Victoria Scahill, and John Lawson. Studies of theory of mind: Are intuitive physics and intuitive psychology independent. *Journal of Developmental and Learning Disorders*, 5(1):47–78, 2001.

Kai Barron, Heike Harmgart, Steffen Huck, Sebastian O Schneider, and Matthias Sutter. Discrimination, narratives, and family history: An experiment with jordanian host and syrian refugee children. *Review of Economics and Statistics*, 105(4):1008–1016, 2023.

Pablo Brañas Garza. Young teens at play: Girls are egalitarian, boys are generous. *Personality and Individual Differences*, 226:12703, 2024.

Pablo Brañas-Garza, Valerio Capraro, and Ericka Rascon-Ramirez. Gender differences in altruism on mechanical turk: Expectations and actual behaviour. *Economics Letters*, 170:19–23, 2018.

Isabelle Brocas and Juan D Carrillo. Studying decision-making in children: challenges and opportunities. *Journal of Economic Behavior & Organization*, 179:777–783, 2020a.

Isabelle Brocas and Juan D Carrillo. The development of social strategic ignorance and other regarding behavior from childhood to adulthood. *Journal of Behavioral and Experimental Economics*, 85:101524, 2020b.

Isabelle Brocas, Juan D Carrillo, T Dalton Combs, and Niree Kodaverdian. The development of consistent decision-making across economic domains. *Games and Economic Behavior*, 116:217–240, 2019.

Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. Responsibility for what? fairness and individual responsibility. *European Economic Review*, 54(3):429–441, 2010.

Alexander W Cappelen, James Konow, Erik Ø Sørensen, and Bertil Tungodden. Just luck: An experimental study of risk-taking and fairness. *The American Economic Review*, 103 (4):1398–1413, 2013.

Alexander W Cappelen, Johanna Mollerstrom, Bjørn-Atle Reme, and Bertil Tungodden. A meritocratic origin of egalitarian behaviour. *The Economic Journal*, 132(646):2101–2117, 2022.

Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3):817–869, 2002.

Daniel L Chen, Martin Schonger, and Chris Wickens. otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016.

Ramon Cobo-Reyes, Jose J Dominguez, Fernando García-Quero, Brit Grosskopf, Juan A Lacomba, Francisco Lagos, Tracy Xiao Liu, and Graeme Pearce. The development of social preferences. *Journal of Economic Behavior & Organization*, 2019.

Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde. Are risk aversion and impatience related to cognitive ability? *The American Economic Review*, 100(3):1238–1260, 2010.

Catherine C Eckel and Philip J Grossman. Are women less selfish than men?: Evidence from dictator experiments. *The Economic Journal*, 108(448):726–735, 1998.

Pierce D Ekstrom and Christopher M Federico. Personality and political preferences over time: Evidence from a multiwave longitudinal study. *Journal of Personality*, 87(2): 398–412, 2019.

Christoph Engel. Dictator games: A meta study. *Experimental Economics*, 14(4):583–610, 2011.

Ernst Fehr and Klaus M Schmidt. The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1:615–691, 2006.

Ernst Fehr, Helen Bernhard, and Bettina Rockenbach. Egalatarianism in young children. *Nature*, 454(28):1079–1084, 2008.

Ernst Fehr, Daniela Glätzle-Rützler, and Matthias Sutter. The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review*, 64:369–383, 2013.

Donald W Fiske. Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology*, 44(3):329, 1949.

Christina Fong. Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics*, 82(2):225–246, 2001.

Adrian Furnham and Mark Fenton-O'Creevy. Personality and political orientation. *Personality and Individual Differences*, 129:88–91, 2018.

David Gill and Victoria Prowse. A structural analysis of disappointment aversion in a real effort competition. *The American Economic Review*, 102(1):469–503, 2012.

Lewis R Goldberg. The structure of phenotypic personality traits. *American Psychologist*, 48(1):26, 1993.

David M Greenberg, Varun Warrier, Ahmad Abu-Akel, Carrie Allison, Krzysztof Z Gajos, Katharina Reinecke, P Jason Rentfrow, Marcin A Radecki, and Simon Baron-Cohen. Sex and age differences in "theory of mind" across 57 countries using the english version of the "reading the mind in the eyes" test. *Proceedings of the National Academy of Sciences*, 120(1):e2022385119, 2023.

Daniel Kahneman, Jack L Knetsch, and Richard H Thaler. Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5(1):193–206, 1991.

Rena A Kirkland, Eric Peterson, Crystal A Baker, Stephanie Miller, and Steven Pulos. Meta-analysis reveals adult female superiority in" reading the mind in the eyes test". *North American Journal of Psychology*, 15(1), 2013.

David A Kolb, Richard E Boyatzis, and Charalampos Mainemelis. Experiential learning theory: Previous research and new directions. In *Perspectives on thinking, learning, and cognitive styles*, pages 227–247. Routledge, 2014.

James Konow. Fair shares: Accountability and cognitive dissonance in allocation decisions. *The American Economic Review*, 90(4):1072–1091, 2000.

John A List. On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3):482–493, 2007.

Erzo FP Luttmer and Monica Singhal. Culture, context, and the taste for redistribution. *American Economic Journal: Economic Policy*, 3(1):157–79, 2011.

Sonny Mattek, Samantha Chavez, Julia L Berkowitz, M Ida Gobbini, and Robert S Chavez. 100 peers: An open set of controlled emotional facial expressions for use in psychology and neuroscience experiments. Unpublished, 2021.

Robert R McCrae and Paul T Costa. Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1): 81, 1987.

Bregtje Gunther Moor, Zdeňa A Op de Macks, Berna Güroğlu, Serge ARB Rombouts, Maurits W Van der Molen, and Eveline A Crone. Neurodevelopmental changes of reading the mind in the eyes. *Social Cognitive and Affective Neuroscience*, 7(1):44, 2012.

Thomas Erik Neuber. Egocentric norm adoption. Technical report, ECONtribute Discussion Paper, 2021.

Hessel Oosterbeek, Randolph Sloof, and Gijs Van De Kuilen. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics*, 7(2):171–188, 2004.

Rike Pahnke, Anett Mau-Moeller, Alfons O Hamm, and Alexander Lischke. Reading the mind in the eyes of children test (rme-ct): Development and validation of a complex emotion recognition test. *Frontiers in Psychiatry*, 11:376, 2020.

Mary E Rogers and A Ian Glendon. Development and initial validation of the five-factor model adolescent personality questionnaire (ffm–apq). *Journal of Personality Assessment*, 100(3):292–304, 2018.

Christopher J Soto and Jennifer L Tackett. Personality traits in childhood and adolescence: Structure, development, and outcomes. *Current Directions in Psychological Science*, 24 (5):358–362, 2015.

Matthias Sutter, Claudia Zoller, and Daniela Glätzle-Rützler. Economic behavior of children and adolescents – a first survey of experimental economics results. *European Economic Review*, 111:98–121, 2019.

Jennifer L Tackett. Evaluating models of the personality–psychopathology relationship in children and adolescents. *Clinical Psychology Review*, 26(5):584–599, 2006.

Anna van der Meulen, Simone Roerig, Doret de Ruyter, Pol van Lier, and Lydia Krabbendam. A comparison of children's ability to read children's and adults' mental states in an adaptation of the reading the mind in the eyes task. *Frontiers in Psychology*, 8:594, 2017.

Michele Vecchione, Harald Schoen, José Luis González Castro, Jan Cieciuch, Vassilis Pavlopoulos, and Gian Vittorio Caprara. Personality correlates of party preference: The big five in five big european countries. *Personality and Individual Differences*, 51 (6):737–742, 2011.

Appendix A. Details of design, procedures and instructions

A1. Instructions of observer and competition tasks (*addition/effort* variant)

[Original instructions in English, later translated to Spanish]

Today, we are going to play a few games with you. In all the games, you will earn points that will be converted into money and placed in your virtual wallet. At the end of the experiment we will send to your school email address an Amazon e-giftcard with the money you made. You will get many points today, so hopefully you will be able to get a nice gift card.

TROPHY GAME

There are two tasks in this game: the "Slider Task" and the "Observer Task." The "Slider Task" is a competition. You will be matched in pairs with another student in the room, but you will not know who this person is (see Figure 8 for the corresponding slides).

[SLIDE 1]

You will see sliders with the cursor in a certain position between 0 and 100, in this example it is at 32. Your job will be to move the cursor exactly to the position 50. Whoever puts more cursors in 50 wins. Let's try one on your screen. This is not for real. It's just for you to know how it works.

[Start SLIDER TRAINING]

Put the cursor at 50. Once you have done it, press NEXT. When the time comes to play the Slider Task, you will see many sliders on your screen. It will look like this.

[SLIDE 2]

You will have 90 seconds to put as many cursors as possible at 50. You will have to work fast to beat the other person.

[SLIDE 3]

At the end of the 90 seconds, if you put more cursors at 50 in your screen than the other student puts in his, you will win the trophy game and receive 3 tokens. if you put fewer cursors than the other student, you will lose and get only 1 token. Each token is worth $1, so you will earn either $1 or $3. Is this task clear?

The other task you are going to play is the "Observer Task". In this task, we give you 4 extra tokens that you have to allocate between a winner and a loser of the slider task.

[SLIDE 4]

This is you [point to the screen]. We give you 4 free tokens.

[SLIDE 5]

All you have to do is split the 4 tokens between the winner and the loser. There is only one rule. You have to give at least one token to each. So, do you want to give 3 tokens to the winner who already has 3 tokens and give 1 to the loser who already has 1 token? Or do you want to give 2 and 2? Or do you want to give 1 to the winner and 3 to the loser? Remember each token is worth $1. Your screen will look like this.

[SLIDE 6]

You allocate tokens by clicking on the dashed tokens. If you change your mind, you can move tokens from one player to another. Once you are happy, press OK. Remember that you will play both the Slider Task and the Observer Task. However, it is very important to understand that you will play with different persons. Let's look at this example.

[SLIDE 7]

You play the Observer Task and allocate tokens to two students who play the Slider Task against each other. You also play the Slider Task but with different people. So, you will earn 1 or 3 tokens with the Slider Task and also receive 1, 2 or 3 tokens from an Observer. Since each token is worth $1, you will get between $2 and $6 dollars just with this game.

OK. The computer will now select the groups.

We will start by playing the "Observer Task" and ask you to tell us how you would like to allocate your tokens between two students in this room who will play the "Slider Task" in just a moment. How many to the winner and how many to the loser.

[Start THIRD PARTY DICTATOR]

Please choose the allocation and press OK.

[Wait for participants to finish]

We will now play the "Slider Task". In a moment, you will see all the sliders. When you are ready press START and try to put as many cursors as possible at 50. At the end, you will see how many cursors you managed to put at 50.

[Start SLIDER TASK]

You will now see a reminder telling you what you chose in the Observer Task. Think about it. If you are happy with the choice you made, click OK. If you have changed your mind, click CHANGE and re-do the allocation. Then press OK.

[Start THIRD PARTY REALLOCATION]

On your screen, you will now see a questionnaire. Please answer the questions. We will tell you at the end of the session, whether you won the Slider Task and how many tokens you got in total.

[Start QUESTIONNAIRE 1 about difficulty and performance]

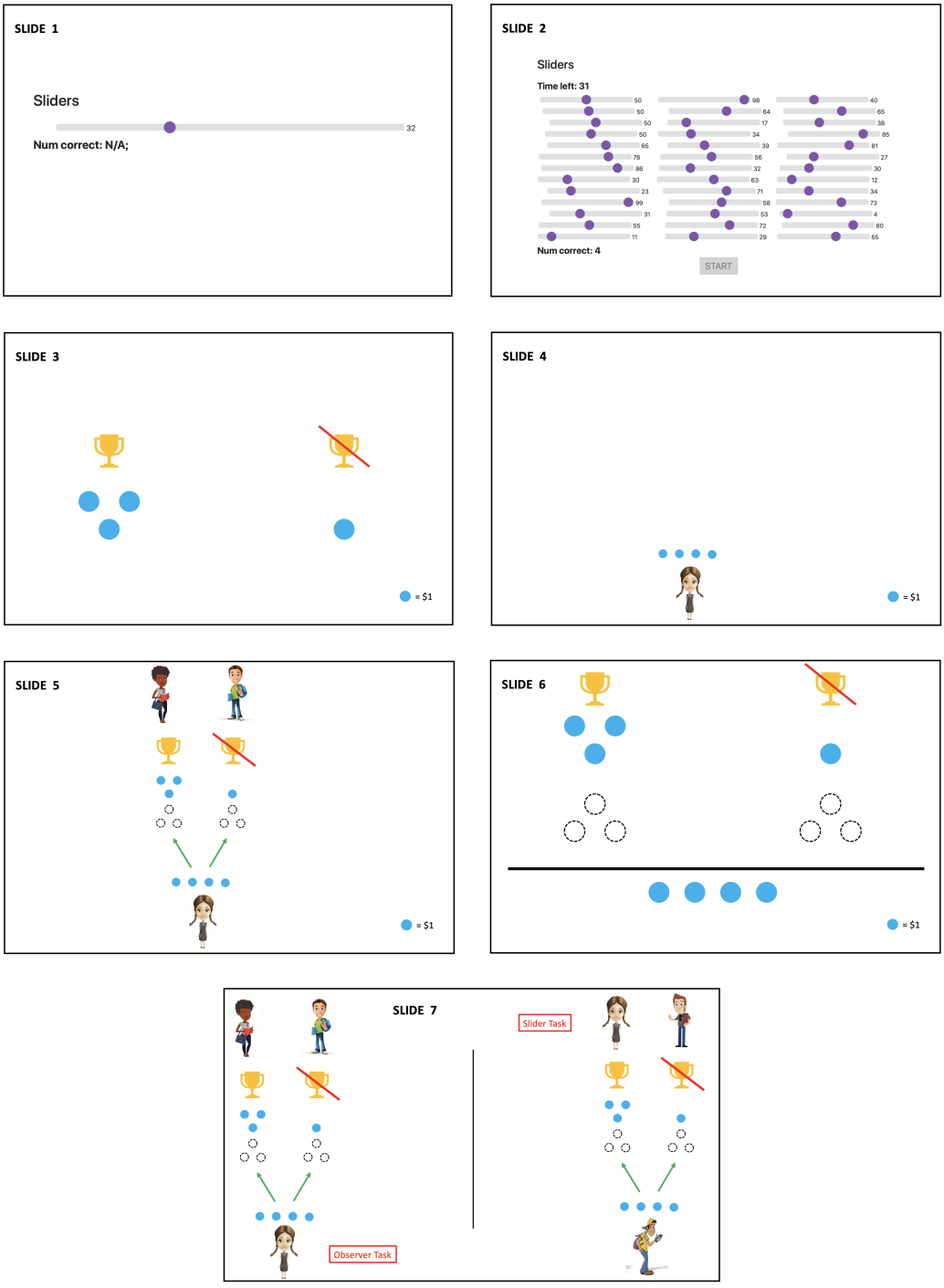


Figure 8: Slides projected on screen for instructions

A2. Developmental Read the Mind in the Eyes Test (DeRMET)

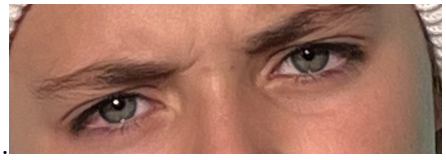
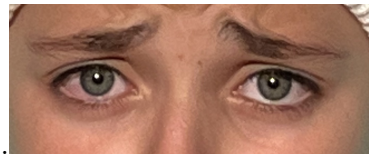
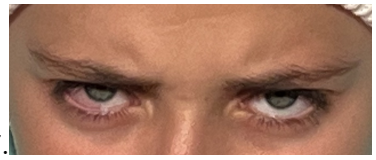
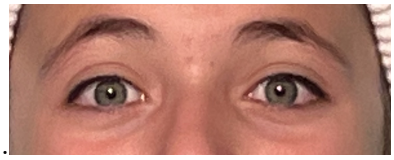
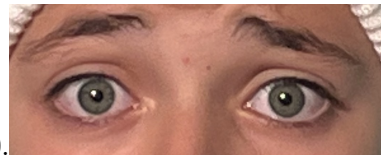
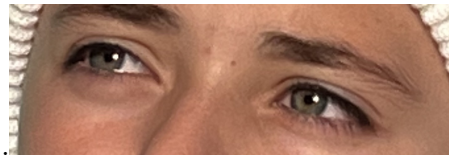
Instructions.

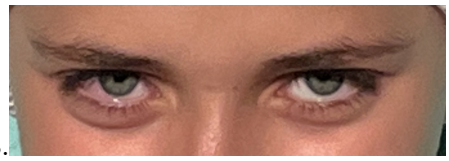
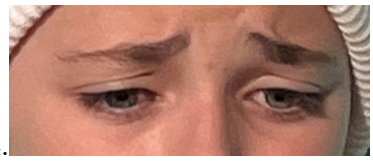
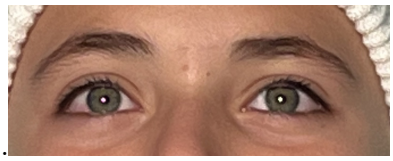
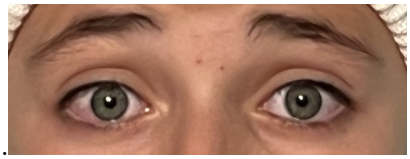
Below, you will see a series of images with a boy's eyes expressing a certain emotion. Look closely and choose, from the four options, the one that best represents the person's emotion. When you are sure of your answer, press NEXT, and the next image will appear. Once you press NEXT, you will not be able to change your answer. For each correct answer, you will receive 10 points, which is equivalent to 10 cents. The more correct answers, the more money you will earn.

Note to the instructor.

Each image is presented in a different page with four emotions to choose from (see the list below). Participant must select exactly one. After a choice is made, the next image appears on the screen and the participant cannot go back. There is no feedback between images.

List of images.

1. 
2. 
3. 
4. 
5. 
6. 
7. 
8. 
9. 
10. 
11. 
12. 

13. 
14. 
15. 
16. 

List of emotions for each image.

Image #	Emotions				% correct
1	1. Bored	2. Confused*	3. Hopeful	4. Guilty	55.2
2	1. Sneaky	2. Angry	3. Guilty	4. Tired*	79.4
3	1. Guilty	2. Tired	3. Confused	4. Sad*	64.7
4	1. Admiring	2. Disapproving*	3. Shocked	4. Bored	45.9
5	1. Surprised	2. Admiring	3. Happy*	4. Sneaky	74.3
6	1. Confused	2. Ashamed*	3. Scared	4. Tired	69.3
7	1. Angry*	2. Scared	3. Sad	4. Worried	97.1
8	1. Excited	2. Confused	3. Thinking	4. Flirting*	67.8
9	1. Excited*	2. Happy	3. Hopeful	4. Flirting	75.3
10	1. Disgusted	2. Sad	3. Angry	4. Scared*	88.2
11	1. Scared	2. Bored	3. Disgusted*	4. Angry	69.8
12	1. Surprised	2. Happy	3. Thinking*	4. Disapproving	86.3
13	1. Hopeful	2. Sneaky*	3. Thinking	4. Excited	86.8
14	1. Guilty	2. Bored	3. Disgusted	4. Worried*	59.7
15	1. Happy	2. Surprised	3. Hopeful*	4. Sneaky	45.9
16	1. Surprised*	2. Happy	3. Flirting	4. Admiring	66.7

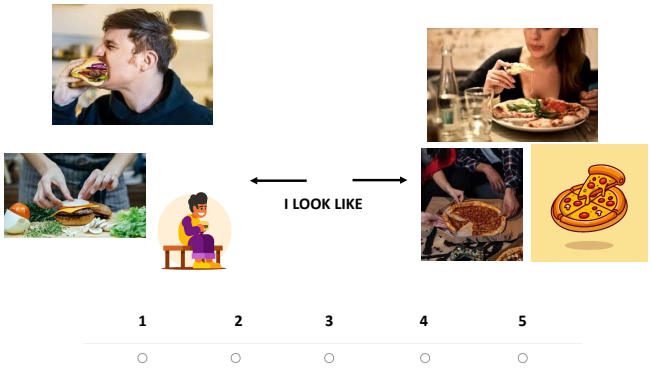
[* = correct answer. The last column reports for each image the aggregate empirical probability of finding the correct answer]

A3. Pictorial Big 5 personality questionnaire

Instructions.

Next, you will see a series of pictures. The group of pictures on the left represents one type of behavior, while the group of pictures on the right represents the opposite type of behavior. Look at the pictures in each group and choose the option that best represents you.

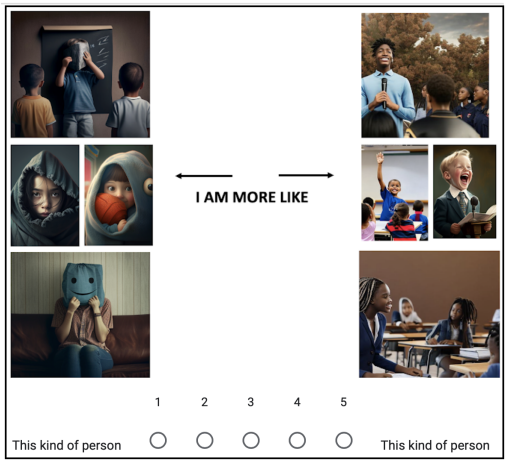
For example, imagine you saw this screen



We would be asking if you definitely prefer hamburgers (1), mostly hamburgers (2), both equally (3), mostly pizza (4) or definitely pizza (5). The questions are about your character or behavior. When you are sure, press NEXT, and another series of pictures will appear. There are no right or wrong answers, and please remember that your responses are confidential.

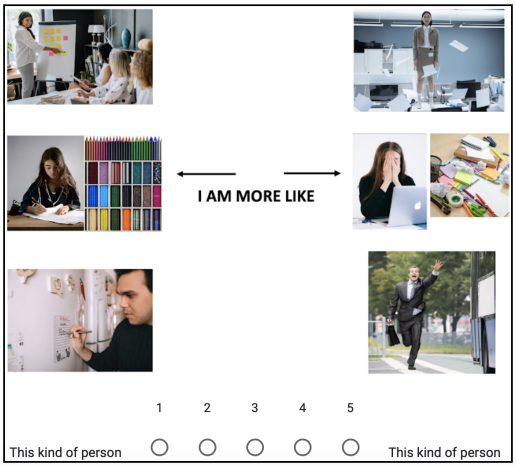
Questionnaire.

The groups of pictures presented to our participants are provided below. They correspond to the following traits: extraversion (1), agreeableness (2), conscientiousness (3), neuroticism (4) and openness (5).

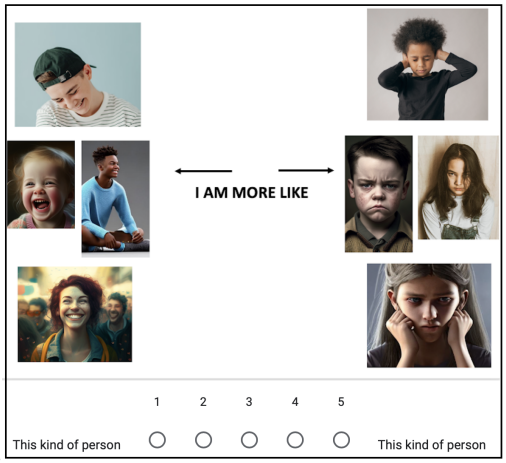
1. 

2. 

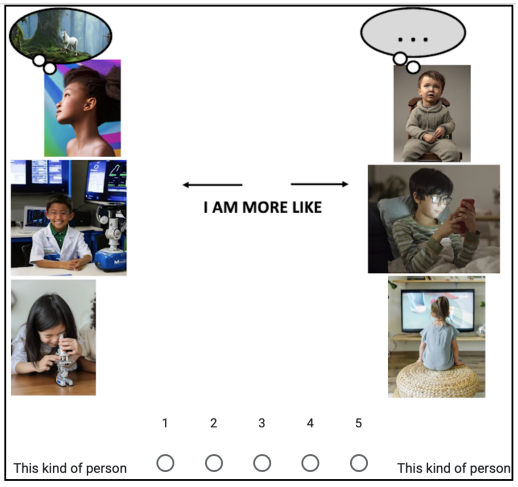
3.


I AM MORE LIKE
1 2 3 4 5
This kind of person
This kind of person

4.


I AM MORE LIKE
1 2 3 4 5
This kind of person
This kind of person

5.


I AM MORE LIKE
1 2 3 4 5
This kind of person
This kind of person

Appendix B. Additional Analyses

B1. Descriptive statistics for the variables of interest

For the descriptive analysis, we cluster students in age-groups according to the Spanish school system: early elementary (US grades 1, 2 and 3), late elementary (US grades 4, 5 and 6), early middle school (US grades 7 and 8), late middle school (US grades 9 and 10) and high school (US grades 11 and 12). Descriptive statistics of the individual characteristics and performance in the tasks are summarized in Table 8.

	Early Elem.	Late Elem.	Early Middle	Late Middle	High School
% Male	47.4	54.6	54.0	52.9	55.7
% Siblings	78.0	88.9	84.1	88.3	86.3
% STEM	40.6	39.6	37.5	47.8	57.0
Pay Winner Before	5.09 (0.81)	5.03 (0.76)	5.05 (0.76)	5.00 (0.75)	5.04 (0.74)
Pay Winner After	5.04 (0.84)	4.97 (0.72)	4.88 (0.71)	4.91 (0.75)	4.95 (0.71)
Score DeRMET	9.12 (2.49)	11.03 (2.43)	12.1 (2.06)	12.56 (1.97)	12.64 (1.75)
Extraversion	3.63 (1.38)	3.74 (1.09)	3.65 (1.05)	3.59 (0.96)	3.50 (0.95)
Agreeableness	4.19 (1.18)	4.09 (1.03)	3.88 (0.94)	3.87 (0.87)	3.90 (0.89)
Conscientiousness	4.06 (1.24)	3.63 (1.27)	3.31 (1.23)	3.03 (1.2)	3.24 (1.19)
Neuroticism	1.59 (1.04)	1.70 (0.96)	1.89 (0.92)	2.03 (0.92)	2.13 (0.93)
Openness	3.43 (1.54)	3.14 (1.3)	2.87 (1.14)	2.85 (1.11)	2.97 (1.05)

Table 8: Summary Statistics by Age Group - mean (standard deviation)

All five groups have a sufficient number of observations and are well-balanced in terms of gender. A good exogenous measure is the percentage of children who have at least one sibling. It appears to be very similar across age groups, with the only difference being in the youngest group (which is precisely the most likely group to increase that number).

Performance in DeRMET increases with age groups up to high school, indicating that this skill develops during adolescence but does not increase beyond the age of 16 (see Appendix B7 for further details). Patterns in the P-Big 5 depend on the trait: agreeableness, conscientiousness and openness decrease with age, neuroticism increases and extraversion remains roughly constant. In most cases, the variables reach their peak (or minimum) in late middle school (see Appendix B8 for further details).

B2. Gender effects

Table 9 reports a further analysis of gender effects by interacting gender with Environment and Condition variables. Male participants allocate less to the loser: the male coefficient is negative and statistically significant for the initial allocation, while the corresponding effect in the final allocation is negative but not statistically different from zero. Task effects are generally small on average, but the interaction patterns indicate some heterogeneity: in the initial decision, the *talent* coefficient in the *addition* environment is negative and marginally significant in the specification with environment interactions, while under *redistribution* the talent effect is more favorable to the loser ($Redistribution \times Talent$ is positive and significant). In the final decision, there is evidence that males respond differently to the *effort* condition: $Effort \times Male$ is negative and statistically

	Initial	Initial	Final	Final
Age	0.003	0.003	0.013^{**}	0.012^{**}
	(0.006)	(0.006)	(0.006)	(0.006)
Age^2	−0.000	−0.000	-0.000^{**}	-0.000^{**}
	(0.000)	(0.000)	(0.000)	(0.000)
Talent	−0.047	$-0.206^{\cdot}$	−0.113	−0.107
	(0.085)	(0.116)	(0.076)	(0.102)
Effort	−0.034	−0.030	0.107	0.188
	(0.084)	(0.108)	(0.083)	(0.114)
Redistribution	-0.188^{***}	-0.302^{**}	-0.167^{**}	−0.111
	(0.070)	(0.118)	(0.065)	(0.123)
Male	-0.128^{*}	-0.150^{*}	−0.103	−0.123
	(0.073)	(0.090)	(0.066)	(0.077)
Siblings	−0.036	−0.038	−0.034	−0.035
	(0.046)	(0.046)	(0.046)	(0.045)
STEM	0.007	0.005	0.023	0.019
	(0.033)	(0.033)	(0.034)	(0.034)
$Redistribution \times Talent$		0.343^{**}		−0.016
		(0.168)		(0.152)
$Redistribution \times Effort$		0.022		−0.156
		(0.163)		(0.168)
$Talent \times Male$	−0.022	0.078	−0.105	−0.064
	(0.086)	(0.129)	(0.080)	(0.112)
$Effort \times Male$	−0.090	−0.112	-0.197^{**}	-0.181^{*}
	(0.086)	(0.116)	(0.083)	(0.096)
$Redistribution \times Male$		0.110		0.132
		(0.123)		(0.117)
$Talent \times Redistribution \times Male$		−0.214		−0.083
		(0.163)		(0.158)
$Effort \times Redistribution \times Male$		0.019		−0.043
		(0.165)		(0.161)
Num. obs.	2267	2267	2267	2267
R^2	0.025	0.030	0.038	0.041
Adj. R^2	0.019	0.022	0.032	0.033

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Campus FE (4 campuses). SE clustered at the session level (131 sessions).

Table 9: Gender effects in the *initial* and *final* allocations to the loser.

significant, implying a larger reduction in the loser's allocation among males in the effort condition, while the remaining two-way and three-way interactions (including $Redistribution \times Male$ and the triple interactions) are not statistically distinguishable from zero.

B3. Replication of the main regressions studying age effects of the *initial* and *final* payoffs of the loser with alternative specifications.

We consider two alternative specifications, both without campus fixed effects. Table 10 and Table 12 cluster standard errors at the session level. Table 11 and Table 13 cluster standard errors at both the session and campus levels.

	Redistribution	Redistribution	Addition	Addition	All	All
Age	−0.001	−0.010	0.002**	0.014·	0.002**	0.014*
	(0.001)	(0.007)	(0.001)	(0.007)	(0.001)	(0.007)
Age^2		0.000		−0.000		−0.000
		(0.000)		(0.000)		(0.000)
Talent	0.063	0.061	−0.149	−0.155*	−0.055	−0.058
	(0.106)	(0.102)	(0.091)	(0.092)	(0.071)	(0.070)
Effort	−0.056	−0.053	−0.070	−0.079	−0.071	−0.073
	(0.099)	(0.099)	(0.094)	(0.090)	(0.068)	(0.067)
Male	−0.114**	−0.108**	−0.160***	−0.161***	−0.140***	−0.138***
	(0.045)	(0.044)	(0.052)	(0.052)	(0.035)	(0.034)
Siblings	−0.055	−0.050	−0.018	−0.026	−0.034	−0.036
	(0.067)	(0.067)	(0.062)	(0.059)	(0.045)	(0.044)
STEM	0.017	0.010	−0.009	0.001	0.005	0.006
	(0.048)	(0.047)	(0.045)	(0.044)	(0.033)	(0.032)
Redistribution					0.277	1.740**
					(0.197)	(0.686)
$Age \times Redistribution$					−0.003*	−0.024**
					(0.001)	(0.010)
$Age^2 \times Redistribution$						0.000**
						(0.000)
Constant	3.062***	3.711***	2.870***	2.081***	2.833***	2.040***
	(0.158)	(0.477)	(0.179)	(0.515)	(0.156)	(0.496)
Num. obs.	1078	1078	1189	1189	2267	2267
R^2	0.010	0.014	0.035	0.042	0.029	0.034
Adj. R^2	0.005	0.007	0.030	0.036	0.025	0.029

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 10: *Initial* allocation to the loser by treatments and age with clustered standard errors at the classroom level.

	Redistribution	Redistribution	Addition	Addition	All	All
Age	−0.001	−0.010	0.002**	0.014*	0.002**	0.014*
	(0.001)	(0.006)	(0.001)	(0.005)	(0.001)	(0.005)
Age^2		0.000		−0.000*		−0.000*
		(0.000)		(0.000)		(0.000)
Talent	0.063	0.061	−0.149	−0.155	−0.055	−0.058
	(0.089)	(0.086)	(0.103)	(0.102)	(0.075)	(0.076)
Effort	−0.056	−0.053	−0.070	−0.079	−0.071	−0.073
	(0.104)	(0.105)	(0.068)	(0.071)	(0.055)	(0.057)
Male	−0.114	−0.108	−0.160**	−0.161**	−0.140**	−0.138**
	(0.052)	(0.050)	(0.037)	(0.037)	(0.032)	(0.031)
Siblings	−0.055	−0.050	−0.018	−0.026	−0.034	−0.036
	(0.067)	(0.062)	(0.053)	(0.052)	(0.049)	(0.047)
STEM	0.017	0.010	−0.009	0.001	0.005	0.006
	(0.028)	(0.019)	(0.021)	(0.017)	(0.017)	(0.014)
Redistribution					0.277	1.740*
					(0.201)	(0.611)
Age × *Redistribution*					−0.003	−0.024*
					(0.001)	(0.010)
Age^2 × *Redistribution*						0.000
						(0.000)
Constant	3.062***	3.711***	2.870***	2.081***	2.833***	2.040**
	(0.145)	(0.361)	(0.069)	(0.330)	(0.100)	(0.366)
Num. obs.	1078	1078	1189	1189	2267	2267
R^2	0.010	0.014	0.035	0.042	0.029	0.034
Adj. R^2	0.005	0.007	0.030	0.036	0.025	0.029

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 11: *Initial* allocation to the loser by treatments and age with clustered standard errors at the classroom and campus levels.

	Redistribution	Redistribution	Addition	Addition	All	All
Age	0.000	−0.002	0.002**	0.022***	0.002**	0.022***
	(0.001)	(0.007)	(0.001)	(0.008)	(0.001)	(0.008)
Age^2		0.000		−0.000***		−0.000***
		(0.000)		(0.000)		(0.000)
Talent	−0.209***	−0.209***	−0.112	−0.122	−0.158***	−0.163***
	(0.078)	(0.077)	(0.090)	(0.085)	(0.060)	(0.058)
Effort	−0.081	−0.080	0.055	0.040	−0.015	−0.022
	(0.097)	(0.096)	(0.096)	(0.087)	(0.068)	(0.065)
Male	−0.113**	−0.111**	−0.204***	−0.205***	−0.158***	−0.158***
	(0.055)	(0.055)	(0.043)	(0.043)	(0.035)	(0.035)
Siblings	−0.016	−0.014	−0.051	−0.064	−0.035	−0.042
	(0.057)	(0.058)	(0.072)	(0.067)	(0.048)	(0.045)
STEM	0.022	0.020	−0.003	0.013	0.011	0.019
	(0.051)	(0.051)	(0.046)	(0.044)	(0.034)	(0.033)
Redistribution					0.172	1.701**
					(0.221)	(0.768)
Age × *Redistribution*					−0.002	−0.024**
					(0.001)	(0.010)
Age^2 × *Redistribution*						0.000**
						(0.000)
Constant	3.095***	3.263***	2.906***	1.568**	2.920***	1.554***
	(0.153)	(0.509)	(0.214)	(0.595)	(0.194)	(0.593)
Num. obs.	1078	1078	1177	1177	2255	2255
R^2	0.018	0.018	0.042	0.061	0.031	0.041
Adj. R^2	0.012	0.012	0.037	0.055	0.027	0.037

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 12: *Final* allocation to the loser by treatments and age with clustered standard errors at the classroom level.

	Redistribution	Redistribution	Addition	Addition	All	All
Age	0.000	−0.002	0.002**	0.022***	0.002**	0.022***
	(0.002)	(0.012)	(0.000)	(0.004)	(0.000)	(0.004)
Age^2		0.000		−0.000***		−0.000**
		(0.000)		(0.000)		(0.000)
Talent	−0.209**	−0.209**	−0.112*	−0.122*	−0.158*	−0.163**
	(0.038)	(0.037)	(0.040)	(0.048)	(0.037)	(0.042)
Effort	−0.081*	−0.080*	0.055	0.040	−0.015	−0.022
	(0.027)	(0.028)	(0.068)	(0.076)	(0.025)	(0.028)
Male	−0.113*	−0.111*	−0.204**	−0.205**	−0.158**	−0.158**
	(0.047)	(0.047)	(0.038)	(0.043)	(0.034)	(0.038)
Siblings	−0.016	−0.014	−0.051	−0.064*	−0.035	−0.042
	(0.068)	(0.067)	(0.027)	(0.024)	(0.037)	(0.036)
STEM	0.022	0.020	−0.003	0.013	0.011	0.019
	(0.048)	(0.041)	(0.035)	(0.025)	(0.023)	(0.020)
Redistribution					0.172	1.701
					(0.335)	(1.239)
Age x Redistribution					−0.002	−0.024
					(0.002)	(0.016)
Age^2 x Redistribution						0.000
						(0.000)
Constant	3.095***	3.263**	2.906***	1.568***	2.920***	1.554***
	(0.321)	(0.971)	(0.056)	(0.236)	(0.053)	(0.254)
Num. obs.	1078	1078	1177	1177	2255	2255
R^2	0.018	0.018	0.042	0.061	0.031	0.041
Adj. R^2	0.012	0.012	0.037	0.055	0.027	0.037

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 13: *Final* allocation to the loser by treatments and age with clustered standard errors at the classroom and campus levels.

B4. Replication of the main regressions using age-groups

In Table 14 and Table 15, we replicate the most relevant results of the study using the age-groups described in Appendix B1 instead of the age in months of our participants. We cluster standard errors at the classroom level and we use campus FE in addition in Table 15 (as in the main text).

The first and second columns study the initial decisions of participants. In *redistribution* (column 1), the early-middle-school coefficient is negative and marginally significant, that is, associated with more meritocratic allocations. This corresponds to the slight, imprecisely estimated curvature in Table 3. In *addition* (column 2), coefficients are positive for both early and late middle schoolers, with the largest point estimate in early middle school. Overall, the age-group estimates are qualitatively consistent with Result 2, while the exact shape of the initial profile remains specification-dependent.

The third and fourth columns examine final decisions. In *redistribution* (column 3), we do not detect an age effect, whereas in *addition* (column 4) the point estimates again describe a concave

	Initial		Final	
	Redistribution	Addition	Redistribution	Addition
Late Elem.	−0.012	0.137	−0.023	0.211*
	(0.104)	(0.093)	(0.102)	(0.107)
Early Middle	−0.221*	0.303**	−0.043	0.353***
	(0.113)	(0.116)	(0.119)	(0.133)
Late Middle	−0.062	0.212**	−0.047	0.277**
	(0.113)	(0.104)	(0.133)	(0.117)
High School	−0.058	0.169	0.064	0.157
	(0.120)	(0.159)	(0.110)	(0.137)
Talent	0.045	−0.146	−0.214***	−0.111
	(0.091)	(0.094)	(0.074)	(0.088)
Effort	−0.052	−0.085	−0.083	0.033
	(0.096)	(0.087)	(0.096)	(0.086)
Male	−0.106**	−0.157***	−0.107*	−0.200***
	(0.043)	(0.052)	(0.055)	(0.042)
Siblings	−0.049	−0.012	−0.009	−0.054
	(0.069)	(0.063)	(0.058)	(0.071)
STEM	0.003	0.006	0.017	0.017
	(0.048)	(0.044)	(0.052)	(0.043)
Constant	3.049***	3.038***	3.173***	3.062***
	(0.101)	(0.116)	(0.109)	(0.150)
Num. obs.	1078	1189	1078	1177
R^2	0.020	0.040	0.020	0.052
Adj. R^2	0.012	0.033	0.011	0.045

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 14: Replication of regressions in Tables 3 and 4 with age-groups (clusters SE at the classroom level)

pattern with a similar peak. Overall, the age-group specifications are qualitatively consistent with the profiles described in Result 3.

	Initial		Final	
	Redistribution	Addition	Redistribution	Addition
Late Elem.	0.002	0.125	0.015	0.209*
	(0.102)	(0.090)	(0.099)	(0.108)
Early Middle School	−0.204*	0.285**	0.009	0.349**
	(0.119)	(0.121)	(0.116)	(0.133)
Late Middle School	−0.053	0.199*	0.006	0.279**
	(0.130)	(0.104)	(0.145)	(0.118)
High School	−0.074	0.194	0.051	0.185
	(0.119)	(0.167)	(0.114)	(0.145)
Talent	0.032	−0.131	−0.218***	−0.095
	(0.086)	(0.089)	(0.074)	(0.085)
Effort	−0.058	−0.076	−0.086	0.048
	(0.095)	(0.088)	(0.091)	(0.089)
Male	−0.105**	−0.156***	−0.104*	−0.199***
	(0.044)	(0.052)	(0.056)	(0.043)
Siblings	−0.049	−0.017	−0.016	−0.058
	(0.070)	(0.063)	(0.058)	(0.070)
STEM	0.006	0.008	0.021	0.020
	(0.047)	(0.043)	(0.051)	(0.043)
Num. obs.	1078	1189	1078	1177
R^2	0.023	0.043	0.027	0.056
Adj. R^2	0.012	0.034	0.016	0.046

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 15: Replication of regressions in Tables 3 and 4 with age-groups (clusters SE at the classroom level and Campus FE)

B5. Final choices in *redistribution* with separate regressions

Table 16 and Figure 9 present the same analysis of final choices in the *redistribution* treatment as Table 4 (columns 1 and 2) and Figure 6a, except that regressions are performed separately for the three winner selections methods. In the main text, we advocated for one regression with dummy variables to increase statistical power. Although the separate regressions rest on about 350 observations per condition and should be read cautiously, it is still instructive to check whether the age trajectory is similar across conditions.

The age trajectory, or rather its absence, is similar when regressions are performed separately. The fitted curves suggest that the youngest participants (age 9 and below) redistribute less to the loser under *luck* than under *effort*, but this comparison rests on few observations and we do not interpret it. The tendency of males to choose more meritocratic allocations is present, at the 10% level, for *effort* and *talent*, and absent when winners are chosen randomly.

B6. Final choices in *addition* with separate regressions

Table 17 and Figure 10 present the same analysis of final choices in the *addition* treatment as Table 4 (columns 3 and 4) and Figure 6b, again with regressions performed on each of the three winner selections methods in order to check the age-trajectory separately for each case.

	effort		*talent*		*luck*	
	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.004	0.005	0.003	0.004	0.004	0.003
	(0.012)	(0.012)	(0.005)	(0.005)	(0.01)	(0.015)
Age^2	−0.00002	0.00002	−0.00000	−0.00000	−0.00001	−0.00001
	(0.00004)	(0.00004)	(0.00002)	(0.00002)	(0.00005)	(0.00005)
Male		−0.185*		−0.147*		0.019
		(0.102)		(0.071)		(0.100)
Siblings		0.023		0.004		−0.068
		(0.094)		(0.112)		(0.094)
STEM		0.011		0.113		−0.027
		(0.074)		(0.097)		(0.101)
Obs.	383	383	333	333	362	362
Adj. R^2	0.018	0.024	-0.006	-0.002	0.006	-0.000

*p<0.1; **p<0.05; ***p<0.01

Table 16: OLS regressions of *final* allocation in the *redistribution* treatment, performed separately on each winner selection method

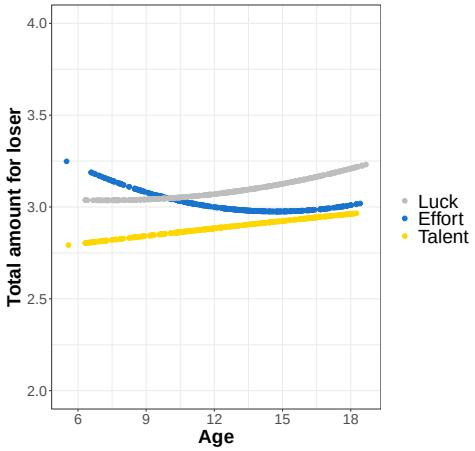


Figure 9: Quadratic best fit in *final* allocation in the *redistribution* treatment, performed separately on each winner selection method

	effort		*talent*		*luck*	
	(3)	(4)	(5)	(6)	(7)	(8)
Age	0.019	0.019*	0.009	0.011	0.041***	0.040***
	(0.012)	(0.011)	(0.013)	(0.012)	(0.009)	(0.009)
Age^2	−0.0001*	−0.0001**	−0.00003	−0.00003	−0.0001***	−0.0001***
	(0.00003)	(0.00003)	(0.00002)	(0.00004)	(0.00003)	(0.00003)
Male		−0.331***		−0.180**		−0.129
		(0.055)		(0.071)		(0.085)
Siblings		−0.038		−0.053		−0.020
		(0.142)		(0.104)		(0.089)
STEM		−0.022		−0.038		0.060
		(0.079)		(0.066)		(0.066)
Obs.	329	329	395	395	453	453
Adj. R^2	0.057	0.094	0.014	0.024	0.130	0.127

*p<0.1; **p<0.05; ***p<0.01

Table 17: OLS regressions of *final* allocation in the *addition* treatment, performed separately on each winner selection method

Again, the point estimates are qualitatively similar when regressions are performed separately, supporting the concave age profile of allocation to the loser in *addition*. Finer grade-by-condition comparisons are less precise and should be interpreted cautiously. Males are estimated to be more meritocratic across the winner-selection methods.

B7. Developmental RMET (DeRMET)

a. Performance

Figure 7 (left) summarizes the distribution of the number of correct answers in the DeRMET task by grade, with the grade average displayed at the top.

Like many tasks that involve a combination of sensory perception and the ability to interpret social cues, performance is not constant across age groups. The improvement is large and significant between grades 1 and 5 (8.1 to 11.3, for an average increase of 0.80 per grade, Pearson Correlation Coefficient between score and age PCC = 0.59, $p < 0.001$). The improvement is then small but still significant between grades 5 and 12 (11.3 to 12.8, for an average increase of 0.21 per grade, PCC = 0.16, $p < 0.001$). It is consistent with an RMET study examining age-related changes of a-ToM in adolescents and adults (Moor et al., 2012). Most importantly, very few participants obtain a perfect score, and the distribution of correct choices is also very far from what we would observe under random guessing for any grade (with random choice, 81% would obtain a score of 5 or less). This suggests that the task is appropriately challenging and suitable for participants across the entire age range of the study. Furthermore, the percentage of correct answers for each image varies between 45.9% and 97.1% (see Appendix A2). This implies that the answer we designated as "correct" is, for every image, the one chosen most often by participants. If this were not the case, it would bring into question what constitutes a correct answer: the one we anticipate or the one that individuals assert. In our experiment, the least clearcut case is image 15, where 46% of participants answer correctly and 32% provide the same wrong answer among the remaining

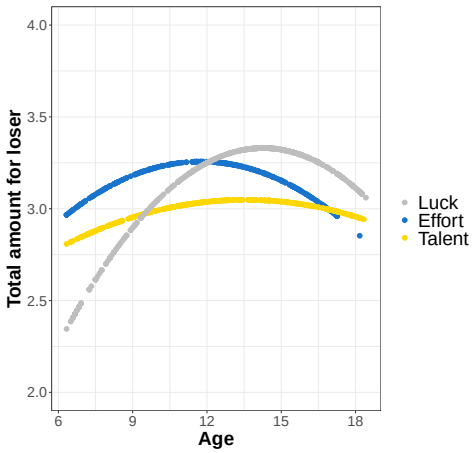


Figure 10: Quadratic best fit in *final* allocation in the *addition* treatment, performed separately on each winner selection method

three. Finally, we find that females are better than males at judging emotions (t-test, $p < 0.001$). This aligns with a study of the RMET conducted over 57 countries (Greenberg et al., 2023) and a meta-analysis of the task (Kirkland et al., 2013).

b. Validation

As briefly mentioned in section 2, we pre-validated the DeRMET task by enlisting 43 teenagers and adults, who were asked to identify emotions in 20 original images. They were also asked to provide feedback on the perceived difficulty of each question. Based on their responses, we removed four emotions that participants found particularly difficulty to identify. The task and validation materials were registered with the Open Science Framework (OSF) at https://osf.io/7kw4m/ to ensure transparency and accessibility.

We report a validation exercise based on the 2,280 observations collected in the experiment reported here. We did not ask respondents to complete the standard RMET designed for children and adolescents. We therefore validate our task independently and from scratch. Additional details are available on the OSF page.

High schoolers answered most questions correctly 70% of the time and above. Exceptions were "sad" (53% of correct answers) and "worried" (61% of correct answers). Taking all participants together, performance varied between 46% and 97% reflecting age-related differences.

Reliability analyses using Cronbach's Alpha (mean 0.67, range 0.60–0.82 from younger to older participants) confirmed internal consistency. Exploratory Factor Analysis (EFA) identified a three-factor structure. Roughly, the first factor represents complex social emotions (e.g., sneaky, flirting, hopeful, disapproving). The second represents fundamental emotion reactivity (e.g., happy, angry, excited) and the third represents sensitivity to negative emotions (e.g., scared, ashamed, disgusted). This structure was valid for both younger (< 13 years old) and older children (confirmatory analysis, Root Mean Square Error of Approximation = 0.022–0.026).

To assess convergent validity, we examined the data of the oldest participants who also completed the mini-International Personality Item Pool (IPIP) questionnaire. We found that DeRMET

scores were significantly associated with items from the "Agreeableness" trait. These items are "I am not really interested in others", "I sympathize with others' feelings", "I am not interested in other people's problems", and "I feel others' emotions". Correlations were small in magnitude–they varied from 0.09 to 0.12–but they all had the expected sign and were strongly significant ($p < 0.001$). This supports the idea that individuals with higher Theory of Mind abilities tend to be more empathetic and considerate. Surprisingly, however, there were no association with extraversion.

Overall, this validation exercise positions DeRMET as a valuable tool for assessing emotional recognition across developmental stages.

B8. Pictorial Big Five (P-Big 5)

a. Performance

Figure 7 (right) presents the average score by grade for each of the five personality traits.

We find significant trends in the reports of all five traits. Most notably, we observe decreasing levels of conscientiousness (PCC between score and age $= -0.27$, $p < 0.001$) and openness (PCC $= -0.16$, $p < 0.001$) and increasing levels of neuroticism (PCC $= 0.20$, $p < 0.001$) as individuals age. Developmental shifts in the expression and maturation of traits have been documented in longitudinal data (Tackett, 2006). Our result is partly consistent with the finding that mean levels of agreeableness, conscientiousness, and openness decline from late childhood to late adolescence (Soto and Tackett, 2015).

b. Validation

Once again, and as mentioned in section 2, we pre-validated the P-Big 5 task by enlisting 40 teenagers and adults, who completed the P-Big 5 as well as a 25-question variant of the Five-Factor Model designed for adolescents. To effectively validate the task, we asked students from grades 5 and up in the experiment reported in this article to complete the mini-IPIP, a 20-question version of the International Personality Item Pool. Below, we present a basic analysis of the relationship between the P-Big 5 and the mini-IPIP (for a more comprehensive validation analysis, please refer to our OSF repository at: https://osf.io/8jbc6/).

To understand the link between these two personality tasks, we regressed P-Big 5 scores on mini-IPIP scores while controlling for age and gender, as these factors may influence how participants respond to visual stimuli. The results are summarized in Table 18.

According to Table 18, there is a very strong association between the two measures, suggesting that both assessments capture similar trait components, even for the case of Neuroticism, which was the least well-captured trait in the pre-validation.

We also ran a convergent validity analysis by examining the relationship between the four individual questions in the mini-IPIP and the single question in the P-Big 5 for each trait. The results are presented in Table 19.

With the exception of Q4, every other individual question of the mini-IPIP was very significantly correlated with the single question of the P-Big5. However, some correlations were stronger than others, indicating that certain facets of the mini-IPIP might be better captured than others.

Overall, our findings suggest that the P-Big 5 ranks individuals similarly to conventional assessments. Because it is significantly shorter, does not require verbal understanding, and can be administered to children (rather than to parents), it is a valuable tool for assessing personality

	Pict O	Pict C	Pict E	Pict A	Pict N
mini-IPIP O	0.436***				
	(0.035)				
mini-IPIP C		0.716***			
		(0.027)			
mini-IPIP E			0.558***		
			(0.025)		
mini-IPIP A				0.251***	
				(0.028)	
mini-IPIP N					0.310***
					(0.029)
Male	−0.144***	−0.193***	0.164***	0.012	−0.126***
	(0.054)	(0.051)	(0.044)	(0.046)	(0.046)
Grade	−0.040***	−0.071***	−0.017*	−0.038***	0.073***
	(0.013)	(0.012)	(0.010)	(0.010)	(0.010)
Constant	1.901***	1.616***	1.767***	3.308***	0.442***
	(0.162)	(0.138)	(0.130)	(0.138)	(0.138)
Adj. R^2	0.094	0.317	0.235	0.054	0.100
Num. obs.	1644	1644	1644	1644	1644

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 18: Regression of P-Big 5 scores on mini-IPIP scores

O - P-Big5		**C - P-Big5**		**E - P-Big5**		**A - P-Big5**		**N - P-Big5**	
Q5	0.22 (p=0)	Q3	0.36 (p=0)	Q1	0.34 (p=0)	Q2*	-0.10 (p=0)	Q4*	-0.04 (p=0.08)
Q8*	-0.12 (p=0)	Q6	0.32 (p=0)	Q7*	-0.32 (p=0)	Q9	0.11 (p=0)	Q10	0.19 (p=0)
Q11*	-0.18 (p=0)	Q13*	-0.37 (p=0)	Q12	0.36 (p=0)	Q14*	-0.18 (p=0)	Q15	0.23 (p=0)
Q17*	-0.21 (p=0)	Q18*	-0.49 (p=0)	Q20*	-0.28 (p=0)	Q16	0.21 (p=0)	Q19*	-0.17 (p=0)

Table 19: Correlations between P-Big 5 scores and IPIP scores by question. * indicates IPIP questions that are framed in the opposite direction of the P-Big 5.

traits in young children, studying developmental trajectories of affective processes, and including it in experiments as a swift and convenient assessment for correlational purposes.

B9. Interactions between traits and Treatment arms

The following tables analyze interactions between traits and treatment in both *initial* and *final* allocations. In the *initial* decision, the environment shift toward *redistribution* increases the likelihood of choosing the meritocratic allocation relative to the equitable benchmark, while the *luck* condition (relative to *effort*) increases the propensity to choose the compensatory allocation. Trait-dependent differences emerge primarily

through a few interactions: for "Merit.", the *Talent* condition is associated with a significantly smaller propensity to choose "Merit." for more extraverted participants (negative $Extraversion \times Talent$), and this pattern is partly offset in *redistribution* (positive three-way $Extraversion \times Talent \times Redistribution$). For "Comp.", the three-way interaction between DeRMET scores and *luck* in *redistribution* is negative, indicating that under *luck* the environment shift is associated with a smaller tendency to choose "Comp." (vs. Equitable) for higher-ToM participants.

In the *final* decision, the environment effect toward *redistribution* remains strong for "Merit.", while the *talent* condition substantially reduces the likelihood of choosing "Comp." Some trait-by-condition interactions persist for "Merit." (notably negative $Extraversion \times Talent$ and $Extraversion \times Luck$), and there is modest evidence that these condition effects may vary with the environment for a small subset of traits (e.g., positive $Extraversion \times Talent \times Redistribution$ and $Conscientiousness \times Luck \times Redistribution$), but overall the three-way interactions are few and do not point to a pervasive pattern.

Table 20: Multinomial regressions of meritocratic and compensatory initial choices: interaction effects

	Initial choice		Intial choice	
	Merit.	Comp.	Merit.	Comp.
Age	-0.346^{***}	-0.317^{**}	-0.343^{***}	-0.317^{**}
	(0.128)	(0.133)	(0.129)	(0.134)
Age^2	0.011^{**}	0.012^{**}	0.011^{**}	0.012^{**}
	(0.005)	(0.005)	(0.005)	(0.005)
Male	0.490^{***}	0.074	0.474^{***}	0.067
	(0.111)	(0.113)	(0.112)	(0.114)
Siblings	-0.022	-0.138	0.002	-0.142
	(0.148)	(0.150)	(0.149)	(0.151)
STEM	-0.011	-0.026	-0.008	-0.043
	(0.108)	(0.112)	(0.109)	(0.113)
DeRMET (resid)	-0.035	0.075	-0.042	0.008
	(0.108)	(0.052)	(0.061)	(0.055)
Extraversion (resid)	0.045	-0.003	0.320^{**}	-0.029
	(0.101)	(0.107)	(0.139)	(0.128)
Agreeableness (resid)	-0.052	0.014	-0.328^{**}	-0.040
	(0.111)	(0.122)	(0.146)	(0.147)
Conscientiousness (resid)	0.039	-0.094	0.161	-0.142
	(0.088)	(0.092)	(0.118)	(0.109)
Neuroticism (resid)	0.050	-0.036	-0.065	-0.099
	(0.116)	(0.128)	(0.149)	(0.147)
Openness (resid)	-0.015	0.144	-0.068	0.123
	(0.086)	(0.091)	(0.113)	(0.110)
redistribution	0.687^{***}	0.094	0.735^{***}	-0.189
	(0.105)	(0.109)	(0.188)	(0.204)

Continued on next page

Table 20 continued

	Initial choice		Initial choice	
	Merit.	Comp.	Merit.	Comp.
luck	0.134	0.494^{***}	0.062	0.414^{**}
	(0.128)	(0.133)	(0.198)	(0.178)
talent	−0.078	0.067	−0.007	−0.281
	(0.129)	(0.138)	(0.193)	(0.191)
Extraversion × *redistribution*	−0.038	−0.003	-0.382^{**}	0.106
	(0.100)	(0.107)	(0.180)	0.193
DeRMET × *redistribution*	0.054	0.0003	0.069	0.056
	(0.047)	(0.050)	(0.083)	(0.094)
Agreeableness × *redistribution*	−0.053	0.056	0.383^{**}	0.207
	(0.112)	0.122)	(0.193)	(0.221)
Conscientiousness × *redistribution*	0.169^{*}	0.181^{**}	−0.029	0.306^{*}
	0.087)	(0.092)	(0.153)	0.164)
Neuroticism × *redistribution*	−0.121	−0.035	0.194	0.122
	(0.118)	(0.127)	(0.207)	(0.240)
Openness × *redistribution*	−0.023	−0.019	0.100	0.005
	(0.086)	(0.090)	(0.149)	(0.163)
DeRMET × *luck*	-0.138^{**}	-0.124^{**}	-0.158^{*}	−0.033
	(0.057)	(0.061)	(0.086)	0.081)
DeRMET × *talent*	−0.026	−0.093	0.010	-0.148^{*}
	(0.058)	(0.063)	(0.085)	(0.086)
Extraversion × *luck*	0.069	0.148^{*}	−0.248	0.282^{*}
	(0.121)	(0.128)	(0.182)	(0.166)
Extraversion × *talent*	-0.271^{**}	-0.281^{*}	-0.700^{***}	0.008
	(0.121)	(0.132)	(0.182)	(0.179)
Agreeableness × *luck*	−0.144	-0.281^{*}	0.304	−0.161
	(0.136)	(0.146)	(0.196)	(0.189)
Agreeableness × *talent*	−0.042	−0.045	0.237	0.017
	(0.136)	(0.152)	(0.198)	(0.203)
Conscientiousness × *luck*	−0.027	0.140	−0.196	0.208
	(0.104)	(0.109)	(0.158)	(0.143)
Conscientiousness × *talent*	−0.108	0.099	-0.278^{*}	0.163
	(0.109)	(0.116)	(0.157)	(0.154)
Neuroticism × *luck*	0.059	0.078	0.256	0.201
	(0.141)	(0.153)	(0.200)	(0.191)
Neuroticism × *talent*	−0.065	0.056	0.076	0.054
	(0.146)	(0.162)	(0.209)	(0.218)
Openness × *luck*	0.086	0.056	0.114	0.132
	(0.104)	(0.110)	(0.154)	(0.145)
Openness × *talent*	0.124	−0.124	0.272^{*}	−0.187
	(0.104)	(0.112)	(0.153)	(0.154)
Redistribution × *luck*			0.154	0.190
			(0.265)	(0.275)
Redistribution × *talent*			−0.125	0.766^{***}
			(0.265)	(0.283)
DeRMET × *Redistribution* × *luck*			−0.011	-0.219^{*}
			(0.117)	(0.124)

Continued on next page

Table 20 continued

	Initial choice		Initial choice	
	Merit.	Comp.	Merit.	Comp.
DeRMET × *Redistribution* × *talent*			−0.055	0.102
			(0.118)	(0.130)
Extraversion × *Redistribution* × *luck*			0.494**	−0.068
			(0.252)	(0.266)
Extraversion × *Redistribution* × *talent*			0.758***	−0.171
			(0.252)	(0.270)
Agreeableness × *Redistribution* × *luck*			−0.841***	−0.413
			(0.281)	(0.303)
Agreeableness × *Redistribution* × *talent*			−0.468*	−0.129
			(0.279)	(0.311)
Conscientiousness × *Redistribution* × *luck*			0.271	−0.166
			(0.214)	(0.221)
Conscientiousness × *Redistribution* × *talent*			0.311	−0.195
			(0.223)	(0.237)
Neuroticism × *Redistribution* × *luck*			−0.412	−0.364
			(0.289)	(0.318)
Neuroticism × *Redistribution* × *talent*			−0.263	−0.042
			(0.297)	(0.333)
Openness × *Redistribution* × *luck*			−0.120	−0.167
			(0.214)	(0.224)
Openness × *Redistribution* × *talent*			−0.297	0.136
			(0.213)	(0.228)
Constant	1.548**	1.428*	1.453*	1.578*
	(0.768)	(0.803)	(0.782)	(0.813)
Obs.	2255	2255	2255	2255
AIC	4,756	4,756	4,754	4,754

*p<0.1; **p<0.05; ***p<0.01.

Table 21: Multinomial regressions of meritocratic and compensatory final choices: interaction effects

	Final choice		Final choice	
	Merit.	Comp.	Merit.	Comp.
Age	−0.839***	−0.412***	−0.833***	−0.415***
	(0.135)	(0.130)	(0.136)	(0.130)
Age^2	0.030***	0.015***	0.030***	0.015***
	(0.005)	(0.005)	(0.005)	(0.005)
Male	0.677***	0.025	0.683***	0.020
	(0.119)	(0.108)	(0.120)	(0.108)
Siblings	−0.017	−0.139	0.009	−0.139
	(0.156)	(0.144)	(0.157)	(0.144)
STEM	−0.203*	−0.126	−0.190	−0.124
	(0.116)	(0.106)	(0.117)	(0.107)

Continued on next page

Table 21 continued

	Final choice		Final choice	
	Merit.	Comp.	Merit.	Comp.
DeRMET (resid)	−0.003	0.119**	−0.051	0.109*
	(0.051)	(0.048)	(0.067)	(0.058)
Extraversion (resid)	0.100	0.011	0.244	0.002
	(0.110)	(0.100)	(0.149)	(0.121)
Agreeableness (resid)	−0.140	−0.015	−0.292*	0.031
	(0.120)	(0.113)	(0.157)	(0.139)
Conscientiousness (resid)	0.081	0.045	0.184	0.026
	(0.096)	(0.086)	(0.128)	(0.104)
Neuroticism (resid)	−0.148	−0.001	−0.207	−0.011
	(0.128)	(0.117)	(0.167)	(0.135)
Openness (resid)	−0.208**	−0.014	−0.205*	−0.055
	(0.094)	(0.085)	(0.124)	(0.104)
redistribution	0.272**	−0.100	0.654***	−0.049
	(0.112)	(0.103)	(0.213)	(0.180)
luck	−0.114	−0.024	0.259	0.026
	(0.140)	(0.122)	(0.214)	(0.169)
talent	−0.039	−0.540***	0.201	−0.495***
	(0.138)	(0.132)	(0.212)	(0.180)
$DeRMET \times redistribution.$	0.004	−0.126***	0.093	−0.106
	(0.050)	(0.047)	(0.091)	(0.084)
$Extraversion \times redistribution$	0.081	−0.046	−0.156	−0.052
	(0.105)	(0.101)	(0.196)	0.173
$Agreeableness \times redistribution$	0.005	0.001	0.224	−0.094
	(0.118)	0.115)	(0.211)	(0.196)
$Conscientiousness \times redistribution$	0.091	0.068	−0.078	0.102
	0.094)	(0.086)	(0.167)	(0.147)
$Neuroticism \times redistribution$	0.148	−0.136	0.235	−0.124
	(0.124)	(0.122)	(0.229)	(0.210)
$Openness \times redistribution$	0.081	0.014	0.097	0.104
	(0.091)	(0.085)	(0.164)	(0.148)
$DeRMET \times luck$	−0.085	−0.098*	−0.044	−0.061
	(0.061)	(0.056)	(0.091)	0.078)
$DeRMET \times talent$	−0.002	−0.044	0.105	−0.062
	(0.062)	(0.061)	(0.090)	(0.083)
$Extraversion \times luck$	−0.218*	0.164	−0.328*	0.188
	(0.131)	(0.119)	(0.190)	(0.162)
$Extraversion \times talent$	−0.384***	−0.058	−0.650***	−0.066
	(0.130)	(0.127)	(0.193)	(0.173)
$Agreeableness \times luck$	−0.025	−0.144	0.212	−0.236
	(0.144)	(0.134)	(0.206)	(0.183)
$Agreeableness \times talent$	−0.018	−0.003	0.108	−0.039
	(0.144)	(0.147)	(0.210)	(0.199)
$Conscientiousness \times luck$	0.044	−0.029	−0.157	−0.059
	(0.114)	(0.100)	(0.167)	(0.138)
$Conscientiousness \times talent$	0.063	0.040	−0.005	0.133
	(0.116)	(0.111)	(0.168)	(0.147)

Continued on next page

Table 21 continued

	Final choice		Final choice	
	Merit.	Comp.	Merit.	Comp.
Neuroticism × *luck*	0.166	0.019	0.216	0.004
	(0.153)	(0.140)	(0.216)	(0.181)
Neuroticism × *talent*	0.197	0.027	0.290	0.062
	(0.155)	(0.157)	(0.226)	(0.207)
Openness × *luck*	0.291***	0.136	0.260	0.242*
	(0.113)	(0.101)	(0.163)	(0.140)
Openness × *talent*	0.184*	−0.086	0.228	−0.079
	(0.111)	(0.108)	(0.164)	(0.147)
Redistribution × *luck*			−0.671**	−0.086
			(0.290)	(0.246)
Redistribution × *talent*			−0.363	−0.050
			(0.285)	(0.266)
DeRMET × *Redistribution* × *luck*			−0.071	−0.068
			(0.125)	(0.112)
DeRMET × *Redistribution* × *talent*			−0.182	0.044
			(0.125)	(0.122)
Extraversion × *Redistribution* × *luck*			0.135	−0.029
			(0.270)	(0.241)
Extraversion × *Redistribution* × *talent*			0.484*	0.048
			(0.266)	(0.258)
Agreeableness × *Redistribution* × *luck*			−0.388	0.201
			(0.294)	(0.272)
Agreeableness × *Redistribution* × *talent*			−0.165	0.098
			(0.293)	(0.298)
Conscientiousness × *Redistribution* × *luck*			0.389*	0.079
			(0.233)	(0.201)
Conscientiousness × *Redistribution* × *talent*			0.076	−0.230
			(0.236)	(0.227)
Neuroticism × *Redistribution* × *luck*			−0.059	0.039
			(0.311)	(0.285)
Neuroticism × *Redistribution* × *talent*			−0.155	−0.069
			(0.315)	(0.320)
Openness × *Redistribution* × *luck*			0.074	−0.234
			(0.231)	(0.205)
Openness × *Redistribution* × *talent*			−0.119	−0.022
			(0.225)	(0.218)
Constant	4.531***	2.658***	4.234***	2.668***
	(0.807)	(0.792)	(0.820)	(0.798)
Obs.	2255	2255	2255	2255
AIC	4,682	4,682	4,713	4,713

*p<0.1; **p<0.05; ***p<0.01.

B10. Changes as a function of traits

	redistribution				*addition*			
	Merit.	Equit. (−1)	Equit. (+1)	Comp.	Merit.	Equit. (−1)	Equit. (+1)	Comp.
Age	−0.034	−0.798	−0.677	0.015	0.011	−0.301	−0.822***	−0.026
	(0.005)	(1.076)	(0.469)	(0.022)	(0.007)	(0.581)	(0.217)	(0.028)
Age^2	0.0002	0.028	0.025	−0.0006	−0.0005	0.009	0.031***	0.0001
	(0.0002)	(0.039)	(0.020)	(0.0008)	(0.0003)	(0.023)	(0.007)	(0.0008)
Male	−0.389***	0.670	0.508	0.351	−0.373*	0.031	0.488	0.345*
	(0.125)	(0.649)	(0.754)	(0.385)	(0.209)	(0.054)	(0.326)	(0.193)
Siblings	0.077	1.275***	−0.115	−0.157	0.123	−0.338	−0.269	0.687***
	(0.090)	(0.533)	(0.820)	(0.205)	(0.128)	(0.411)	(0.342)	(0.160)
STEM	0.193	−0.565	−0.653	0.001	0.142	0.001	0.706***	−0.266
	(0.186)	(0.411)	(0.518)	(0.208)	(0.164)	(0.158)	(0.102)	(0.230)
Talent	−0.078	−1.203***	−0.375	0.695***	−0.625**	0.037	0.120	0.115
	(0.211)	(0.469)	(0.579)	(0.155)	(0.262)	(0.249)	(0.428)	(0.307)
Better	−0.434**	−0.961	0.789***	0.633**	−0.872***	−0.180***	0.713***	−0.301
	(0.183)	(0.842)	(0.176)	(0.274)	(0.256)	(0.074)	(0.172)	(0.269)
Worse	0.434***	0.624	−0.548	−1.584*	0.406***	1.412***	−14.276***	−0.910*
	(0.164)	(0.449)	(1.044)	(0.808)	(0.129)	(0.167)	(0.000)	(0.501)
Hard	0.341***	−1.448***	−0.466	1.108***	−0.150	0.884	−0.924**	−0.482
	(0.085)	(0.583)	(0.694)	(0.121)	(0.256)	(0.866)	(0.557)	(0.777)
Easy	0.043	−0.749**	0.248	0.125	−0.190	−0.026	0.306	0.658***
	(0.219)	(0.406)	(0.176)	(0.482)	(0.136)	(0.415)	(0.404)	(0.055)
DeRMET	−0.050	0.030	−0.030	0.071*	0.011	0.096	0.086***	−0.072
	(0.031)	(0.030)	(0.086)	(0.040)	(0.044)	(0.068)	(0.037)	(0.107)
Extraversion	−0.044**	−0.125	0.100	0.156	0.067	−0.089	−0.152	−0.156
	(0.019)	(0.157)	(0.135)	(0.103)	(0.096)	(0.072)	(0.305)	(0.109)
Agreeableness	−0.079	−0.159	0.156	0.0006	0.113	0.012	−0.169***	−0.523*
	(0.099)	(0.203)	(0.179)	(0.101)	(0.121)	(0.168)	(0.085)	(0.081)
Conscienciousness	0.001	0.277	0.150	0.056	−0.175*	0.097	0.153***	0.232
	(0.061)	(0.251)	(0.237)	(0.168)	(0.097)	(0.134)	(0.057)	(0.175)
Neuroticism	−0.087	−0.227	−0.057	0.322***	0.003	0.062	−0.006	−0.223
	(0.067)	(0.154)	(0.324)	(0.111)	(0.078)	(0.121)	(0.270)	(0.171)
Openness	0.075	−0.175	0.017	0.140	0.132	0.115**	−0.136	0.072
	(0.046)	(0.124)	(0.068)	(0.105)	(0.123)	(0.068)	(0.209)	(0.046)
Constant	−0.198	4.405	3.171	−4.473**	−1.076	0.143	2.801	2.074
	(0.460)	(7.081)	(4.988)	(2.273)	(0.620)	(2.523)	(2.679)	(1.445)
Obs.	266	272	272	166	191	353	353	177
AIC	348	590	590	260	519	457	457	321

*p<0.1; **p<0.05; ***p<0.01

Table 22: Poisson and multinomial regressions of change in allocation to winner as a function of initial choice (meritocratic, equitable, compensatory) by treatment with traits as explanatory variables.