

# One size fits all? The interplay of incentives, effort provision, and personality\*

Zvonimir Bašić <sup>†</sup>      Stefania Bortolotti <sup>‡</sup>      Daniel Salicath <sup>§</sup>  
Stefan Schmidt <sup>¶</sup>      Sebastian O. Schneider <sup>¶</sup>      Matthias Sutter <sup>¶,||</sup>

This version: April 23, 2024

## Abstract

Incentives are supposed to increase effort, yet individuals react differently to incentives. We examine this heterogeneity by investigating how personal characteristics, preferences, and socio-economic background relate to incentives and performance in a real effort task. We analyze the performance of 1,933 high-school students under a Fixed, Variable, or Tournament payment. Ability and beliefs about relative performance play a decisive role for productivity when payments are exogenously imposed. Yet, when given the choice to select the payment, also personality traits, economic preferences and socio-economic background matter. Algorithmic assignment of payments could improve productivity, as we show.

**Keywords:** Effort, productivity, incentives, personality traits, preferences, socio-economic background, ability, heterogeneity, sorting, algorithm, lab-in-the-field experiment

**JEL classification:** C93, D91, J24, J41

---

\*This study has been pre-registered at the AEA RCT Registry under AEARCTR-0008360. For helpful discussions and comments we thank Alexander Cappelen, Uri Gneezy and Bertil Tungodden as well as seminar participants at the 8th M-BEPS and the meeting of the European Economic Association in Barcelona. Corresponding author: matthias.sutter@coll.mpg.de

<sup>†</sup>Adam Smith Business School, University of Glasgow, 2 Discovery Place, Glasgow G11 6EY, UK

<sup>‡</sup>University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy

<sup>§</sup>Norwegian Labour and Welfare Administration, Fyrstikkalléen 1, 0661 Oslo, Norway

<sup>¶</sup>Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Straße 10, 53113 Bonn, Germany

<sup>||</sup>University of Cologne, Albertus-Magnus-Platz, 50935 Cologne, Germany; University of Innsbruck, Universitätsstraße 15, 6020 Innsbruck, Austria; University of Venice Ca' Foscari, Canareggio 873, 30100 Venice, Italy.

# 1 Introduction

Understanding how to motivate people to provide effort is of key importance for success in many domains of life, ranging from the educational sector to the labor market. Social scientists and practitioners have long debated on how to best align the interests of principals and agents, and the design of optimal compensation contracts has played a prominent role in this discussion. Extensive theoretical and empirical research has focused on the role of both intrinsic and extrinsic incentives to promote effort provision (Prendergast, 1999; Gneezy et al., 2011; Gneezy and Rey-Biel, 2014; Gneezy et al., 2019; Cassar and Meier, 2018). Yet, it is still relatively poorly understood which incentives are best for which people, as humans react in very different ways to the same incentives. Therefore, it has become a major issue in management to understand the heterogeneity of effort provision in reaction to different incentives (Opitz et al., forthcoming).

In this paper, we examine how a broad set of personal characteristics, skills and preferences, as well as one’s socio-economic background shape performance under various payments. Some people thrive and express their best potential in competitive environments, while some instead choke under such pressure (Dohmen, 2008; Ariely et al., 2009). Some people are diligent and work hard regardless of the environment, while others need monetary rewards to be motivated. Scientific evidence offers surprisingly little guidance in understanding the underpinning of this heterogeneity.

A better understanding of what motivates individuals to put in strong effort under different payments would have far-reaching implications for practitioners as well as for theoretical models. However, field data on personal characteristics and productivity under different incentive contracts are scarce, either because companies do not have data on personality traits or, if they do, it is proprietary to combine these data with information about payments and productivity. Even in the latter case, a proper identification strategy is difficult to achieve because of self-selection as well as market frictions. It is likely that individuals choose different career paths conditional on the incentives offered in the different paths. On top of that, once working under specific incentives, this experience may affect a person’s reaction to different payments. For instance, a competitive environment might shape how one reacts to a tournament payment. In order to mitigate all of these issues, we conducted a controlled large-scale lab-in-the-field experiment in a setting with very limited self-selection and where participants have hardly any workplace experience with different payments.

We ran our experiment in German high schools, with students from grade ten and older (sixteen to twenty years of age, average of 17.1 years). Relying on a pool of high school students offers a number of advantages, which are crucial for our research question. First, they are highly heterogeneous in terms of socio-economic background, traits, and ability levels. This is an advantage of our sample, as different personal characteristics could play a role in their performance, but also their preference for a particular payment. Other sam-

ples – like university students or workers from a particular industry – would be much less heterogeneous than a high school sample from Germany where a large majority of a birth cohort attends high school.<sup>1</sup> Second, high school students have usually not yet experienced different compensation schemes through experiences in the labor market; rather, they are all exposed to similar incentives in schools. As the mentioned experience may shape preferences, this raises a concern about recruiting individuals who already sorted into jobs. Such potential concerns are void in our study. Third, students are close to entering the labor market – either as full-time workers after high school or as part-time workers during tertiary education after high school. This means that our sample, while not yet exposed to (substantial) labor market experience, will soon work under different incentives in labor markets.

In our experiment, we implemented a tedious counting task adapted from Abeler et al. (2011) to measure effort provision. We used three different payments, called “Fixed”, “Variable”, and “Tournament”. The Fixed payment pays a flat wage for doing the real effort task. The Variable payment offers a piece rate per correctly solved task, and the Tournament payment offers a higher piece rate than in “Variable” if a subject performed better than another person, but a lower piece rate otherwise. While certainly other variants of payments could be analyzed (see Opitz et al., forthcoming, for some others), our three payments cover a large share of contracts actually offered on labor markets. We implemented two treatments. In one treatment, we exogenously assigned participants to one of three different payments. In the other treatment, we let participants themselves choose which payment they preferred. The goal of this treatment is threefold. First, it allows us to test if participants are able to self-select into the payment that maximizes their potential monetary earnings. Second, it serves as a test to see if having agency over the type of payment itself has an effect on one’s performance. Third, we can examine whether the same personal characteristics are relevant for sorting into different payments and for being productive with a given payment, or whether the two aspects – choice of payment, and productivity in given payment – are driven by different traits and characteristics.

With regards to personal characteristics and traits, we consider a plethora of factors that have been proven important in understanding labor market outcomes, such as socio-economic status (see, e.g., Heckman, 2006, 2007), personality (see Donato et al., 2017), grit (see, e.g., Duckworth et al., 2007; Alan et al., 2019), competitiveness (see, e.g., Buser et al., 2014), economic preferences, and parenting styles (see, e.g., Bonin et al., 2007; Borghans et al., 2008b; Cadena and Keys, 2015; Reuben et al., 2017; Falk et al., 2018; Kosse and Tincani, 2020; Falk et al., 2023). While the link between some personal characteristics and (labor market) outcomes is quite well established, others are far less understood. Most im-

---

<sup>1</sup>In 2020, 70% of 17-year-olds in Germany were enrolled in some form of upper secondary education; see, e.g., [https://stats.oecd.org/Index.aspx?DataSetCode=EAG\\_ENRL\\_RATE\\_AGE#](https://stats.oecd.org/Index.aspx?DataSetCode=EAG_ENRL_RATE_AGE#).

portantly, it is not a priori clear if and how the above characteristics interact with specific payments.

Our results show that ability in the task and one's own assessment of relative performance are the main drivers of output under all three payments. Personality traits, economic preferences and socio-economic background have at best a marginal influence, which basically confirms that incentives do work, by and large, independently of those personal characteristics. This is not the case, however, as far as sorting is concerned. When subjects can choose among the three payments, personality traits, preferences and socio-economic background matter on top of ability and expectations about own performance. Extraversion and neuroticism, competitiveness, risk and time preferences are predictive of what kind of payment a person chooses. Moreover, sorting does not seem to mainly be focused on optimizing performance by choosing according to the characteristics that are important when assigned to a payment exogenously. Rather, we can show that about half of our subjects would benefit both in earnings and utility from the task if an algorithm was applied to assign them to a particular payment, rather than them having the choice.

Our study makes three main contributions. First, we address unanswered questions on heterogeneity in effort provision. While the literature on the interaction between payments and people's characteristics is still scarce, a notable exception is Donato et al. (2017). In the domain of health care provision, they report that people with high conscientiousness (as one of the Big-5 personality traits) provide better maternal and child services, but react less to performance incentivization. People with low conscientiousness and neuroticism perform well with performance incentivization. Moreover, in a lab experiment, Segal (2012) finds a similar pattern (albeit only for men) between conscientiousness and reaction to incentives. The direction of how the characteristics interact with incentives is for the most part ambiguous as the literature is nascent. We present a systematic account of a variety of traits, preferences and socio-demographic characteristics and how they work to motivate performance with different payments.

Second, we present an extensive analysis of sorting decisions across three different payments. While the previous literature has usually been limited to studying the sorting decisions between two payments (Niederle and Vesterlund, 2007; Buser et al., 2014; Almås et al., 2016), our comprehensive setup allows us to investigate sorting decisions in much greater detail. In addition, we include a wide range of socio-demographics, traits, and preferences that have been found on their own to influence sorting, but we can also examine whether those factors have the same influence both on sorting as well as performance under a specific payment.

Finally, we contribute by investigating which factors determine performance. Do participants understand their performance potential and how it might change across payments? We find that different characteristics matter in determining performance based on whether one is assigned to or choosing the payment. Having the choice therefore plays an important role in assessing which personal characteristics are important for higher performance. However, participants are not mainly sorting based on the characteristics that are influential when assigned to a payment. This points to other factors being prioritized when choosing a payment. It is the latter aspect that is a key distinction and contribution of our paper in comparison to the most closely related paper that is by Opitz et al. (forthcoming). They ran an experiment on MTurk and studied which personal characteristics were the main drivers for effort in a real effort task under six different, exogenously implemented payments. From this main experiment they can estimate the factors that are related to higher performance, and with these estimates they then let a machine learning algorithm assign a new set of MTurkers to the most promising payment, conditional on the new workers' personality traits. The algorithmic assignment increases performance significantly above the level of the single best payment. While we can also estimate how much algorithmic assignment could improve performance – and also utility – our *Endogenous* treatment allows for three further contributions compared to Opitz et al. (forthcoming): first it reveals additional insights into the drivers of choosing a particular payment; second it shows that choices of payments are driven by partly different traits than performance under a given payment; and third it shows that subjects fail in maximizing a particular objective (be it utility or performance or earnings) when given the choice between payments.

The remainder of the paper is organized as follows. In section 2 we present our experimental design. Section 3 reports the results. Section 4 provides insights into how an algorithmic assignment to payments could improve performance, payments and utility of participants. Finally, section 5 concludes the paper.

## 2 Experimental Design

### 2.1 Sample

The experiment was conducted with adolescents in schools across North Rhine-Westphalia, Germany (see Appendix A.1 for a map of participating schools). Altogether 1,914 high school students, enrolled in tenth to thirteenth grade, were recruited and attended both required sessions. Summary statistics of the students in our study are presented in Table 1 (details on the variables and measures are explained in the remainder of this section). We targeted what in German is referred to as “Gesamtschule”: schools that commonly comprise both low and high education tracks within the same institution. This ensured

a heterogeneous sample in our study (with respect to SES, cognitive abilities, etc.). We contacted in a random order the 201 closest schools in the areas of Bonn, Cologne, and Düsseldorf within the state of North Rhine-Westphalia.<sup>2</sup> We first informed and invited schools to participate in the study via a letter. In case of no reply, we contacted the school via phone and sent a more detailed description of the study via email. For every participating school, the study was approved by school principals.<sup>3</sup> Parents were informed about the experiment and needed to sign a consent form in order for a student to participate in the study.<sup>4</sup> Participation was voluntary and it was explicitly mentioned to participants that they could quit the study (or skip specific parts) at any time. As Riener et al. (2020) document the absence of self-selection of schools into experiments in North Rhine-Westphalia, which is where we conducted our experiment, we are confident that our sample is representative of the population of schools in this federal state.

## 2.2 Real Effort Task (RET)

We implemented a counting task adapted from Abeler et al. (2011). Subjects were presented with a sequence of tables containing zeros and ones (Figure 1). The task consists of highlighting and counting the ones present in each table (for instructions, see Section A.5). A table is correctly solved if: (i) all the ones are highlighted, (ii) none of the zeros are highlighted, and (iii) the total amount of ones is correctly reported. Subjects had a total of three trials to solve each table. The task has several desirable features: it does not require any prior knowledge, performance is easy to objectively measure, learning plays only a minor role, and performing the task has no value outside the experiment (Abeler et al., 2011; Charness et al., 2018). Moreover, the task is tedious and requires effort to be solved.

## 2.3 Timeline and Treatments

The experiment comprises two parts conducted approximately two weeks apart from each other (see Table 2 for an overview over the two parts).<sup>5</sup> The same subjects participated in both parts of the experiment. In part 1, a broad range of socio-demographic characteristics, traits, and preferences, as well as a measure of individual productivity was collected from the subject pool. In part 2, effort provision was measured under three different payments: Fixed, Variable, and Tournament payment. Two between-subjects treatments were

---

<sup>2</sup>Contact information is publicly available online on the webpage of the Ministry for School and Education of North Rhine-Westphalia: <https://www.schulministerium.nrw.de/BiPo/SchuleSuchen/online>.

<sup>3</sup>The study has been approved by the ethical board of the University of Innsbruck (certificate of good standing N.o 07/2019 - 25.01.2019).

<sup>4</sup>Students that are 18 or older could sign the consent form themselves.

<sup>5</sup>On a few occasions, the two parts were moved closer/further apart due to logistic reasons. In one case, the two parts took place on two consecutive days; in two other cases, the interval was 5 and 6 weeks, respectively. Overall, the median time between first and second part of the study was 14 days (Mean (SD) of 12.9 (7.6)).

	Mean	SD	N
Productivity part 1 (RET 5 min)	26.94	6.24	1933
Productivity part 2 (RET 20 min)	122.24	22.73	1933
Female (=1)	0.53	0.50	1874
Age	17.12	1.16	1933
Grade (9-13)	11.29	1.00	1933
IQ (Raven 0-10)	5.08	1.48	1898
Grade Math	2.81	1.16	1933
Grade German	2.73	0.91	1933
Born Germany (=1)	0.93	0.25	1933
Speak German at home (=1)	0.94	0.24	1933
Parents German (=1)	0.62	0.49	1933
Mother univ. diploma (=1)	0.14	0.34	1933
Father univ. diploma (=1)	0.18	0.39	1933
One parent univ. diploma (=1)	0.24	0.43	1933
Single parent (=1)	0.21	0.41	1933
Number of siblings	1.64	1.12	1933
Books at home (1-6)	2.36	1.32	1933
Pocket money (0-95)	28.54	26.26	1933
Number of cars (0-3)	1.86	0.84	1933
Number of holidays (0-3)	1.79	1.05	1933
PISA wealth index (0-17)	12.83	2.37	1933
FAS index (0-10)	6.68	1.97	1933
Low SES (=1)	0.22	0.41	1933
Patience (1-32)	18.14	11.98	1933
Patience survey; 0-10)	7.22	1.98	1933
Risk Taking (1-32)	10.22	6.58	1933
Risk Taking (survey; 0-10)	5.92	1.96	1933
Altruism (0-10)	7.49	2.22	1933
Extraversion (cont; 1-5)	3.48	0.78	1933
Agreeableness (cont; 1-5)	3.56	0.58	1933
Conscientiousness (cont; 1-5)	3.35	0.40	1933
Neuroticism (cont; 1-5)	2.90	0.72	1933
Openness (cont; 1-5)	3.43	0.65	1933
Enjoy competition (cont; 1-5)	3.11	1.00	1933
Positive parenting (cont; 1-5)	3.39	0.91	1933
Grit (cont; 1-5)	3.23	0.45	1933
Effort part 2 (1-7)	5.05	1.58	1933
Stress part 2 (1-7)	3.86	1.72	1933
Exhaustion part 2 (1-7)	3.45	1.64	1933
Belief on rel. performance (0-1)	0.55	0.25	1933
Overconfidence	0.06	0.30	1933

Note: We define all variables in full detail in Section A.4, but give here a short description. Note that 59 observations are missing for Female, as some students did not want to reveal their gender. 35 observations are missing for IQ due to technical issues during a session. Books at home is based on six categories for the number of books available at home (0-10, 11-25, 26-100, 101-200, 201-500 and more than 500). The PISA wealth index is based on the family wealth possessions index from the PISA test. The Family Affluence Scale (FAS) is an index to measure family SES based on questions about owning computers, cars, number of vacations, etc. Low SES is a binary variable capturing educational and time resources available to the family. Effort, Stress, and Exhaustion are self-reported measures relating to subjects' experience of our 20-minute real effort task (RET) in part 2 of the study. Belief on rel. performance (0 - 1) is a normalized measure of the subject's belief about own rank in the performance distribution in the RET in part 1 of the study. As session size differs across observations, we normalize the belief on own rank by the total number of participants in the respective session. The measure is, thus, defined between 0 (subject believes to be on the lowest percentile in the distribution) and 1 (subject believes to be on the highest percentile in the distribution). Overconfidence represents the difference between the normalized belief about own relative performance and normalized actual own relative performance. A positive value represents an overconfident self-assessment regarding relative performance in terms of percentiles of the performance distribution.

Table 1: Summary statistics

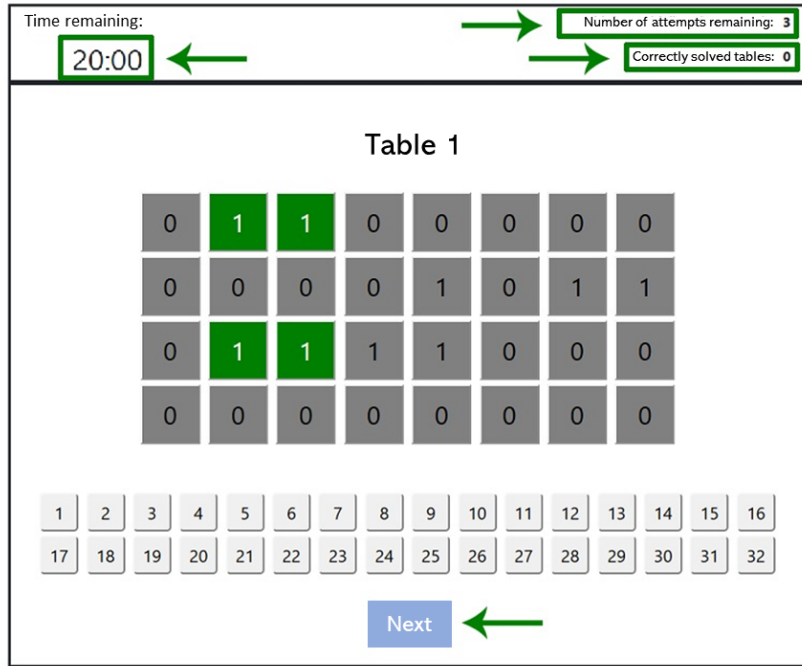


Figure 1: Real Effort Task (RET) – In the center of the screen, participants see a table measuring  $4 \times 8$ , where each square can either display a one or a zero. The objective is to click on every square containing the number one to highlight it. Once this is done, participants should provide the total count of highlighted ones by selecting the corresponding number in the lower white cells. To finalize the task, participants need to click on the “Next” button. If all the ones are correctly highlighted, and the accurate count is submitted, a new table will appear. Otherwise, participants have up to three chances to revise their inputs.

implemented, in which we either imposed a specific payment (*Exogenous* treatment) or subjects could choose their preferred payment (*Endogenous* treatment). In the remainder of this section, we describe in detail the payments, the treatments, the sequence of tasks in the two parts of the experiment, and the data that we collected.

**Payments** We implemented the following three types of payments:<sup>6</sup>

- Fixed payment: a flat payment (€6.5) independent of the number of correctly solved tables;
- Variable payment: subjects were paid a piece rate (€0.06) per correctly solved table;
- Tournament payment: subjects were paid either a high (€0.08) or a low (€0.04) piece rate per correctly solved table. Each participant in this treatment was matched with another participant that also chose the Tournament payment and was paid the high (low) rate if they solved more (less) tables than the matched participant.<sup>7</sup>

<sup>6</sup>Payment amounts for the payments were calibrated based on pilot data to ensure comparable payoffs across payments.

<sup>7</sup>Ties were solved by a random draw of the computer and participants were made aware of that.



<b>Part 1 – Socio-demographics, traits and preferences</b>	<i>Exogenous</i> treatment	<i>Endogenous</i> treatment
Personal ID	✓	✓
RET instructions	✓	✓
Ability (RET 5 min)	✓	✓
IQ test (Raven’s matrices, 5min)	✓	✓
SES questionnaire	✓	✓
Big Five (BFI-44)	✓	✓
Competitiveness (14-item)	✓	✓
Preference module	✓	✓
Positive parenting (6-item)	✓	✓
Grit (8-item)	✓	✓
Average payment (€)	€4 + RET (5 mins)	€4 + RET (5 mins)
Average time	45 min	45 min

<b>Part 2 – Effort provision and payments</b>	<i>Exogenous</i> treatment	<i>Endogenous</i> treatment
Personal ID	✓	✓
Belief elicitation	✓	✓
RET instructions	✓	✓
Instructions (payments)	1 payment (within session randomization) <sup>§</sup>	all 3 payments
Choice (payment)	-	✓
RET (20 min)	✓	✓
Average payment (€)	€1 + RET (20 min) + belief	€1 + RET (20 min) + belief
Average time	45 min	45 min
Number of Observations	944	973

**Notes:** <sup>§</sup>About 1/3 of participants were assigned to each of the three payments.

Table 2: Timeline and overview of the experimental tasks and design

**Treatments** We ran two between-subjects treatments: *Exogenous* and *Endogenous*. In the *Exogenous* treatment, participants were assigned to either the Fixed, the Variable, or the Tournament payment. Subjects only received information about the relevant payment they were assigned to, and were paid accordingly. In the *Endogenous* treatment, participants received information about all the three types of payments and had to choose one payment which determined how their payoff was calculated.

## 2.4 Part 1 – Socio-Demographic Characteristics, Traits and Preferences

The first part of the study was common to all treatments and measured a number of socio-demographic characteristics, traits, and preferences of the participants. We focused on four main areas: ability, family background, preferences, and personality traits (for a detailed list of included questions, see the questionnaire in Section A.5). We started by collecting a measure of ability for the real effort task (RET), where we follow Dohmen and Falk (2011) in incentivizing the task. Participants were given five minutes to solve as many

tables as they could and were paid on a piece rate (€0.06) basis.<sup>8</sup> From this, we create a residualized productivity measure of performance in part 1 to be used as predictor for performance in part 2, along all other predictors. We regress, first, our full set of predictors on performance. Then we use the residuals of this regression as a measure of productivity that is corrected for the correlation between performance in part 1 and the remaining set of our predictors.

After the five-minute RET to assess productivity, a five-minute computerized version of a standardized non-verbal intelligence test was administered (matrix task; Raven, 2000). In addition, we collected demographic information, as well as information about socio-economic status (SES). Our items are informed by three different socio-economic indices, where we also added some own questions. With respect to psychological measures, we collect the Big Five (John and Srivastava, 1999), competitiveness (Newby and Klein, 2014), positive parenting style (Frick, 1991; Essau et al., 2006), and grit (Duckworth and Quinn, 2009). For all these measures, we rely on widely used psychological scales. Finally, we included a series of non-incentivized questions taken from the validated preference module by Falk et al. (2018, 2023) to measure patience, willingness to take risk, and altruism.<sup>9</sup> Given that we elicited fifteen variables that capture socio-economic status, we rely on principal component analysis (PCA) in order to best utilize the extensive data we obtained: Using the weights on the first component resulting from the PCA, we construct a single-item socio-economic status measure including all of the items of three different socio-economic status indices and additional relevant variables. Section A.4 contains a detailed description as well as the motivation behind the selection of included variables and details on the PCA).

## 2.5 Part 2 – Effort Provision and Payments

The second part of the study captured effort provision under the different payments. The RET was the same as in part 1, but lasted for 20 minutes.<sup>10</sup> Before the RET, we elicited participants' beliefs about their performance in part 1. More precisely, we ranked all the participants present in the room based on their performance in part 1 and then asked them to guess their rank.<sup>11</sup> If the guess was correct, they earned €2. If the difference between the guess and the actual ranking was at most 5 positions, they still earned €0.50. Only participants who were present in both visits were included in the ranking (and were asked to guess). We opted for collecting the guesses about their ranking in the second visit.

---

<sup>8</sup>To familiarize themselves with the task and the software, subjects were asked to solve a trial table before moving to the actual task.

<sup>9</sup>For patience and willingness to take risk, we include both qualitative and quantitative items (see the instructions and the questionnaire in Section A.5).

<sup>10</sup>The distribution of number of solved tables in part 1 and part 2 can be found in Figure A2.

<sup>11</sup>On average there were 34.2 participants per session.

We did this as we wanted to examine how beliefs about rankings relate to the choice of payment and to other variables collected in the second session. As students could update their beliefs between the first and the second session, we eliminated this issue by eliciting the beliefs at the beginning of the second session. Feedback about the guessing task was given only at the end of the study.

The beliefs elicited in this way are used in two measures: ‘Belief on rel. performance (perc.)’ is a normalized measure of the subject’s belief about own rank in the performance distribution in the RET in part 1 of the study. As session size differs across observations, we normalize the belief on own rank by the total number of participants in the respective session. The measure is, thus, defined between 0 (subject believes to be on the lowest percentile in the distribution) and 1 (subject believes to be on the highest percentile in the distribution). ‘Overconfidence’, the second measure, represents the difference between the normalized belief about own relative performance and normalized actual own relative performance. A positive value represents an overconfident self-assessment regarding relative performance in terms of percentiles of the performance distribution.

## 2.6 Assignment to treatments

The assignment of subjects into treatments happened on the session-level. In Table 3, we show that samples assigned to either of the two treatments, *Exogenous* and *Endogenous*, are comparable and do not differ along any of the elicited characteristics used in later analyses.

In the *Exogenous* treatment we randomly assigned participants within a session to one of the three payments based on the min MSE method developed by Schneider and Schlather (2017). Based on re-randomization, the method aims at minimizing the mean squared error of the treatment effect estimator as a function of treatment assignment. The method thus increases precision of the treatment effect estimation by choice of treatment assignment. Intuitively, the method forms comparable treatment groups considering multivariate information such as gender, SES, ability, etc. We opt for this method, as it allows us to assign three treatments in the same session while still “balancing” multivariate and continuous information in a principled way.<sup>12</sup> Moreover, balance with respect to the considered variables is less affected in case of attrition (Schneider and Schlather, 2017). To achieve balanced treatment groups, we consider pre-treatment information on the productivity in the RET, demographic information, socio-economic characteristics, psychological measures as well as preferences. The method was applied at the school level. In Table 4, we investigate whether our explanatory variables across different payments in *Exogenous* are balanced by testing whether at least one group is different from the other two groups using Kruskal-Wallis tests. Of the 36 comparisons, none is significant not even at the 10%

---

<sup>12</sup>Previous work mostly relies on binary assignment between two different treatments.

level, indicating that our treatment assignment was successful in creating well balanced groups.

In the *Endogenous* treatment, students were first introduced to the three available payments and were then allowed to pick their most preferred one before starting to work for 20 minutes on the RET.

Characteristic	Overall	Treatments		p-value
		Exogenous	Endogenous	
<b>Skills</b>				
Grade German	2.73 (0.91)	2.72 (0.92)	2.74 (0.91)	0.77
Grade Math	2.81 (1.16)	2.82 (1.13)	2.79 (1.19)	0.50
IQ (Raven 0-10)	5.08 (1.46)	5.13 (1.43)	5.02 (1.49)	0.088
Productivity part 1 (RET 5 min)	26.94 (6.24)	26.84 (6.19)	27.03 (6.29)	0.53
<b>Demographics</b>				
Age rel. to grade mean	-0.01 (0.81)	0.01 (0.77)	-0.03 (0.84)	0.11
Female (=1)	0.53 (0.49)	0.53 (0.49)	0.54 (0.49)	0.72
Grade (9-13)	11.29 (1.00)	11.28 (0.97)	11.29 (1.02)	0.91
Number of siblings	1.64 (1.12)	1.68 (1.13)	1.59 (1.11)	0.068
Positive parenting (1-5)	3.39 (0.91)	3.40 (0.91)	3.38 (0.91)	0.49
Composite SES Index	0.00 (1.00)	-0.03 (0.99)	0.03 (1.01)	0.25
<b>Personality Traits, Economic Preferences, and Beliefs</b>				
Altruism (0-10)	7.49 (2.22)	7.53 (2.20)	7.46 (2.25)	0.57
Extraversion (1-5)	3.48 (0.78)	3.47 (0.78)	3.48 (0.78)	0.95
Agreeableness (1-5)	3.56 (0.58)	3.56 (0.57)	3.57 (0.59)	0.46
Conscientiousness (1-5)	3.35 (0.40)	3.36 (0.42)	3.35 (0.39)	0.66
Neuroticism (1-5)	2.90 (0.72)	2.91 (0.73)	2.88 (0.71)	0.42
Openness (1-5)	3.43 (0.65)	3.41 (0.66)	3.45 (0.65)	0.24
Enjoy competition (1-5)	3.11 (1.00)	3.09 (0.98)	3.12 (1.01)	0.56
Grit (1-5)	3.23 (0.45)	3.23 (0.46)	3.22 (0.45)	0.41
Belief on rel. performance (0 - 1)	0.55 (0.25)	0.55 (0.24)	0.54 (0.25)	0.35
Patience Index	-0.01 (0.80)	0.00 (0.79)	-0.01 (0.82)	0.94
Risk Index	0.01 (0.75)	0.00 (0.75)	0.02 (0.76)	0.53
<b>Number of Observations</b>	<b>1933</b>	<b>983</b>	<b>950</b>	

Note: The p-values report results from Wilcoxon rank-sum tests of differences between the two treatment groups.

Table 3: Balance in treatment assignment for *Exogenous* and *Endogenous*

## 2.7 Procedures

To avoid self-selection into the study, it was conducted in schools during regular school hours. Sessions were run in large lecture halls and several classes took part in the experiment at the same time.<sup>13</sup> The number of participants in a single session was on average 34.5 with a 12.9 standard deviation. The experiment was conducted with up to 75 tablets and a server using oTree (Chen et al., 2016).

<sup>13</sup>Due to logistic constraints, some sessions were conducted in single classrooms.

Characteristic	Overall	Incentive Schemes			p-value
		Fixed	Variable	Tournament	
<b>Skills</b>					
Grade German	2.72 (0.92)	2.71 (0.96)	2.72 (0.88)	2.72 (0.91)	0.94
Grade Math	2.82 (1.13)	2.81 (1.07)	2.83 (1.18)	2.82 (1.13)	>0.99
IQ (Raven 0-10)	5.13 (1.43)	5.14 (1.37)	5.16 (1.45)	5.09 (1.49)	0.84
Productivity part 1 (RET 5 min)	26.84 (6.19)	26.91 (6.25)	26.77 (6.52)	26.83 (5.79)	0.91
<b>Demographics</b>					
Age rel. to grade mean	0.01 (0.77)	0.03 (0.78)	-0.04 (0.75)	0.02 (0.78)	0.48
Female (=1)	0.53 (0.49)	0.52 (0.50)	0.53 (0.49)	0.53 (0.49)	0.94
Grade (9-13)	11.28 (0.97)	11.32 (0.98)	11.27 (0.95)	11.25 (0.99)	0.60
Number of siblings	1.68 (1.13)	1.62 (1.09)	1.71 (1.12)	1.73 (1.17)	0.49
Positive parenting (1-5)	3.40 (0.91)	3.38 (0.91)	3.40 (0.91)	3.42 (0.92)	0.79
Composite SES Index	-0.03 (0.99)	0.01 (1.03)	-0.05 (0.93)	-0.04 (1.01)	0.75
<b>Personality Traits, Economic Preferences, and Beliefs</b>					
Altruism (0-10)	7.53 (2.20)	7.47 (2.27)	7.54 (2.20)	7.57 (2.12)	0.94
Extraversion (1-5)	3.47 (0.78)	3.47 (0.77)	3.46 (0.78)	3.50 (0.80)	0.80
Agreeableness (1-5)	3.56 (0.57)	3.54 (0.56)	3.55 (0.57)	3.58 (0.57)	0.85
Conscientiousness (1-5)	3.36 (0.42)	3.34 (0.42)	3.35 (0.40)	3.37 (0.44)	0.76
Neuroticism (1-5)	2.91 (0.73)	2.95 (0.76)	2.92 (0.70)	2.88 (0.73)	0.56
Openness (1-5)	3.41 (0.66)	3.39 (0.68)	3.44 (0.60)	3.40 (0.68)	0.62
Enjoy competition (1-5)	3.09 (0.98)	3.08 (1.00)	3.13 (0.95)	3.07 (0.99)	0.73
Grit (1-5)	3.23 (0.46)	3.23 (0.50)	3.24 (0.46)	3.24 (0.42)	0.81
Belief on rel. performance (0-1)	0.55 (0.24)	0.54 (0.25)	0.56 (0.24)	0.56 (0.24)	0.31
Patience Index	0.00 (0.79)	0.04 (0.78)	-0.02 (0.80)	-0.02 (0.78)	0.65
Risk Index	0.00 (0.75)	0.00 (0.77)	0.00 (0.73)	0.00 (0.74)	0.90
<b>Number of Observations</b>	<b>983</b>	<b>331</b>	<b>327</b>	<b>325</b>	

Note: The p-values report results from Kruskal-Wallis tests of whether at least one group is different from the other groups.

Table 4: Balance for Payments in *Exogenous*

In each of the two parts, subjects were randomly assigned to a desk upon arrival. They were all separated by privacy screens, and communication was strictly forbidden throughout the experiment. This was enforced to avoid students comparing choices or their performance. Teachers were allowed to be in the classroom but were not allowed to communicate with or observe the behavior of the participants. In the first part, the relevant instructions were read aloud, and displayed on the screens before the beginning of the RET and IQ task. In the second part, subjects were reading the instructions displayed on the screen alone, since multiple payments were randomized within the same experimental session in the *Exogenous* treatment. To ensure that subjects fully understood the payments, they had to individually answer a set of computerized control questions before proceeding with the task itself.

Since subjects took part in two separate parts, data was matched via a personal ID created by the participants at the beginning of each part (see the instructions in Section A.5). Each part lasted around 45 minutes (a regular school hour) and participants were paid anonymously and in cash. In part 1, participants received a fixed payment of €4, plus the earnings for the 5 minutes RET. In part 2, participants earned a €1 show-up fee, plus

the earnings for the 20 minutes RET and from the guessing task (beliefs). On average, participants earned €5.65 in part 1 and €8.71 in part 2, which is in total roughly in the range of what is recommended as weekly allowance for that age group.<sup>14</sup>

We aimed for a sample size of about 2000, as a result of analytical and simulation-based power calculations.<sup>15</sup> Data was collected between March 2019 and August 2022.<sup>16</sup>

## 3 Results

We first discuss which factors are predictive of performance in the different payments when they are exogenously assigned, before turning to the results in the case participants themselves have the choice between payments. We conclude with investigating subjects' sorting decisions, and then present estimations how performance, earnings, and utility from the task could be improved through an algorithmic assignment of participants to payments on the basis of personality traits and background characteristics.

### 3.1 Heterogeneity in Effort Provision – *Exogenous* treatment

The light blue bars in Figure 2 illustrate performance across payments in the *Exogenous* treatment. Performance is measured as the total number of correctly solved tables in the 20-minute real effort task. We see that average performance is in the range from 120 to 125 correct tasks. Despite the relatively small range, we see significant differences across payments, as participants with the Variable payment performed, on average, better than participants with the Fixed payment ( $t$ -test,  $p < 0.01$ ) and those with the Tournament payment ( $p < 0.05$ ).

Next we investigate this heterogeneity in effort provision by examining how they are related to individual characteristics. The results are reported in Table 5. We have structured the table into three groups of variables, with skills at the top, demographics in the middle and personality traits, economic preferences and beliefs at the bottom. In the upper panel, we see that math grades and productivity (measured as the performance in the 5-minute RET in part 1 of the experiment) have a positive influence on performance (in part 2). IQ is unrelated to performance. In the middle panel, we note that older partici-

---

<sup>14</sup>See, e.g., <https://www.dji.de/themen/jugend/taschengeld.html> (in German; last accessed: 12/07/2023).

<sup>15</sup>With this sample size, we have 80% power for detecting interaction effects of moderate size (i.e., 30% of a main effect), see our pre-registration for details and formulae.

<sup>16</sup>Data collection was paused several times during the COVID-19 pandemic because of school closures. Although we had pre-registered to end data collection in May 2022, for some schools, it was not possible to offer us a date for conducting sessions before, even though we had contacted them at the beginning of the school year, i.e., September of the previous year. We accepted their offer for later dates nonetheless, thereby maintaining a good relationship with schools and openness towards future research requests by us and others in our region.

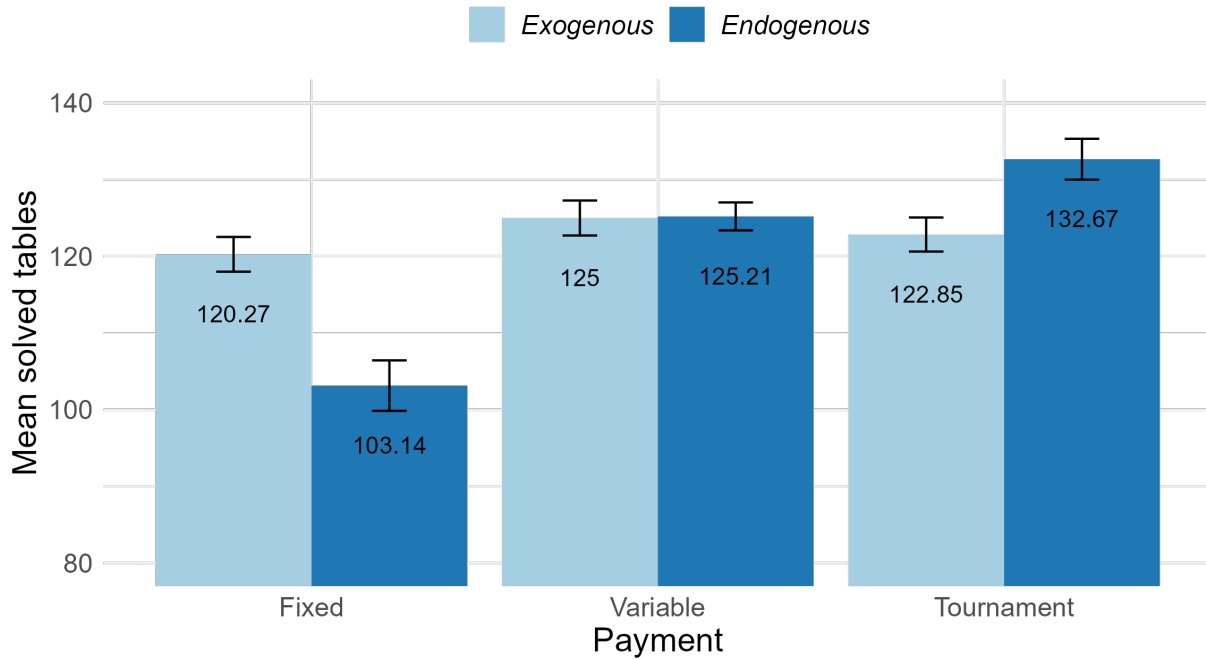


Figure 2: Mean completed tasks in part 2, by treatment and payment in part 2

*Notes:* Whiskers (in black) indicate 95% confidence intervals. In the *Exogenous* treatment, there is a significant difference in performance, where those participants assigned to the Variable payment performed better than both, those assigned to the Fixed payment as well as those assigned to the Tournament payment (paired  $t$ -test of difference:  $p < 0.01$  and  $p < 0.05$ , respectively). There are no differences between the Fixed and the Tournament payments. In the *Endogenous* treatment, we observe a significant difference in performance in all pairwise comparisons of payment conditions, where those in the Fixed payment performed the poorest, followed by those in the Variable payment, with participants in the Tournament payment performing best (paired  $t$ -test in all pairwise comparisons:  $p < 0.01$ ).

pants (recall they are between 16 and 20 years old) are better performing, while those who are relatively older in their respective grade perform worse. The latter might likely be due to participants who had to repeat a grade (which happens in about 10% to 20% of cases in Germany) and are thus relatively older (and given the repetition of a grade on average less able or motivated than others). Interestingly, socio-economic status is not related to performance. We see some effects of gender (women performing better in Fixed) and positive parenting (participants with more caring parents perform worse in the Tournament payment), but these effects are not found across all payments. In the bottom panel, it is noteworthy that not a single personality trait of the Big-5 has a significant influence on performance, nor are economic preferences (risk and time preferences, competitiveness) or grit relevant. Only one's belief about the relative performance is significant. This variable is scaled from 0 to 1, spanning the range from expecting to be the worst performer (0) to expecting to be the best performer (1). The coefficients of around 24 indicate that someone is estimated to solve 24 more tasks if the person believes to be the best performer rather than the worst performer. In other words, if someone believes to move up one decile in

relative performance, the person solves about 2.4 tasks more.

Overall, it seems that there are hardly any differences regarding the predictive quality of the individual characteristics across payments, and a single regression with interaction terms confirms this (see Table ?? in Appendix A.2). Sizeable differences worth highlighting concern the math grade (highest and positive correlation for Fixed, lowest, insignificantly positive correlation in the case of the Tournament payment), and positive parenting, which leads to worse results with the Tournament payment. More generally, the results in Table 5 suggest that personality traits and SES are not very influential for performance under the different payments.<sup>17</sup>

	Incentive Schemes		
	Fixed	Variable	Tournament
	(I)	(II)	(III)
<b>Skills</b>			
Grade German	-0.523 (0.935)	-0.339 (1.765)	0.537 (1.152)
Grade Math	4.319 (0.904)***	2.162 (0.846)**	1.265 (0.924)
IQ (Raven 0-10)	0.152 (0.532)	0.418 (0.648)	0.814 (0.585)
Productivity (resid.)	1.704 (0.223)***	1.568 (0.273)***	1.883 (0.315)***
<b>Demographics</b>			
Age rel. to grade mean	-2.697 (0.791)***	-3.615 (1.257)***	-1.907 (1.225)
Female (=1)	4.725 (2.561)*	2.687 (2.492)	1.723 (2.319)
Grade (9-13)	4.243 (0.808)***	3.625 (1.579)**	4.440 (1.039)***
Number of siblings	-0.266 (1.309)	-0.964 (0.835)	0.912 (0.774)
Positive Parenting (1-5)	0.881 (1.382)	-0.952 (1.109)	-2.225 (1.076)**
Composite SES Index	-0.203 (0.861)	-0.651 (1.023)	-1.048 (1.138)
<b>Personality Traits, Economic Preferences, and Beliefs</b>			
Altruism (0-10)	-0.303 (0.420)	0.729 (0.473)	-0.012 (0.585)
Extraversion (1-5)	-1.643 (1.291)	-0.009 (1.225)	-0.701 (1.569)
Agreeableness (1-5)	1.321 (2.833)	-1.416 (1.947)	-1.934 (1.890)
Conscientiousness (1-5)	0.082 (2.386)	0.736 (4.171)	1.986 (2.743)
Neuroticism (1-5)	-0.353 (1.258)	-0.605 (1.343)	-1.201 (1.563)
Openness (1-5)	1.354 (1.524)	-0.361 (1.923)	0.968 (1.419)
Enjoy Competition (1-5)	0.083 (1.311)	1.622 (1.281)	1.978 (1.301)
Grit (1-5)	-1.968 (1.831)	-0.636 (2.425)	-1.600 (1.999)
Belief on rel. performance (0-1)	23.438 (4.511)***	24.198 (4.593)***	22.552 (4.107)***
Patience Index	1.212 (1.099)	0.562 (1.229)	-0.609 (1.488)
Risk Taking Index	1.697 (1.177)	-0.979 (1.197)	-1.939 (1.695)
Constant	43.498 (21.244)**	61.658 (26.858)**	55.851 (19.719)***
Num.Obs.	331	327	325
R2 Adj.	0.343	0.342	0.341

Note: Table shows OLS regressions of performance on characteristics in part 2 split by treatment and incentive scheme. Standard errors (in parentheses) clustered on the session level. Productivity (resid.) is a residualized measure of performance in part 1. For this, we regress our full set of predictors on performance. We use the residuals of this regression as a measure of productivity that is corrected for the correlation between performance in Part 1 and the remaining set of our predictors.

Table 5: Productivity by Payment – *Exogenous* treatment

<sup>17</sup>Results are robust to using LASSO regressions, see Table ?? in Appendix A.2.



### 3.2 Heterogeneity in Effort Provision – *Endogenous* Treatment

The dark blue bars in Figure 2 show large differences in the average performance when payments have been chosen by participants themselves, with the output in Tournament being almost 30% higher than in Fixed. Comparing the light blue bars (for the *Exogenous* treatment) with the dark blue bars (for the *Endogenous* treatment), we note that self-selection matters a lot. In Fixed, performance drops markedly by about 20%, while in Tournament it increases by about 8% in the *Endogenous* treatment. This already hints at selection effects, and we are going to study the factors for selection in the next subsection.

In Table 6 we present the regression results on which factors are related to performance with a given payment. This table is identically structured as the previous Table 5. Comparing both tables, we note that also in the *Endogenous* treatment, two variables are robustly related to performance, which are (the residualized) productivity in part 1 and beliefs about one’s relative performance. While here the patterns are very similar across both treatments, Table 6 also reveals differences. Age (both absolute as the grade one is in, and relative compared to the grade mean) plays practically no role any longer. The bottom panel reveals that personality traits and economic preferences become more important in the *Endogenous* treatment, which may not be so surprising, given that participants can make their own choice about the payment, which is related to personal characteristics, as we will see later. Regarding personality traits, we observe that higher conscientiousness is associated with worse performance when choosing the Tournament payment, while the reverse is true (albeit insignificant) when choosing the Variable or the Fixed payment. Moreover, in line with Donato et al. (2017), we find a negative interaction between conscientiousness and tournament payments, i.e., a significant difference in coefficients of conscientiousness in Tournament compared to both, the Variable and the Fixed payment (see Table ?? in Appendix A.2). We also see effects (at the 10% significance level) of extraversion in case of the Fixed payment (again with significant differences to the other payments, see Table Table ??), and of agreeableness for the Variable payment. Enjoying competition makes participants more productive in Variable, but surprisingly not in Tournament (although there is no significant difference between the two coefficients).

The most noteworthy significant difference in terms of predictive quality of a variable is probably productivity: In the Fixed payment, its predictive quality is blurred compared to the Variable and Tournament payment, where coefficients are about twice as large.<sup>18</sup>

Overall, compared to the *Exogenous* treatment, the evidence from the *Endogenous* treatment suggests that performance in case of having agency over the payment is partly driven by other factors than when the payment is exogenously assigned.

---

<sup>18</sup>See Tables ?? and ?? in Appendix A.2 for regressions including interaction terms, from which statements about significance regarding differences in coefficients between the payments are inferred.

	Incentive Schemes		
	Fixed	Variable	Tournament
	(I)	(II)	(III)
<b>Skills</b>			
Grade German	-3.301 (1.509)**	-0.506 (0.948)	1.703 (1.232)
Grade Math	1.871 (1.391)	1.958 (0.745)***	0.967 (1.088)
IQ (Raven 0-10)	1.064 (1.451)	0.783 (0.491)	1.293 (0.876)
Productivity (resid.)	1.253 (0.579)**	2.501 (0.225)***	2.294 (0.301)***
<b>Demographics</b>			
Age rel. to grade mean	0.973 (1.363)	-0.366 (0.978)	-0.543 (1.115)
Female (=1)	-0.180 (5.274)	4.173 (1.553)***	3.133 (3.559)
Grade (9-13)	1.251 (1.979)	1.794 (0.834)**	2.399 (1.519)
Number of siblings	-1.211 (1.099)	0.012 (0.708)	0.680 (1.389)
Positive Parenting (1-5)	-2.635 (2.221)	-1.118 (0.680)	-0.310 (1.120)
Composite SES Index	0.512 (1.401)	-0.911 (0.485)*	1.582 (1.216)
<b>Personality Traits, Economic Preferences, and Beliefs</b>			
Altruism (0-10)	0.977 (1.048)	0.300 (0.368)	-0.553 (0.417)
Extraversion (1-5)	-4.048 (2.220)*	-0.068 (1.387)	1.721 (1.242)
Agreeableness (1-5)	-2.951 (3.887)	2.687 (1.524)*	2.232 (1.467)
Conscientiousness (1-5)	6.322 (5.400)	2.944 (2.267)	-6.592 (3.284)**
Neuroticism (1-5)	-0.945 (3.584)	0.516 (1.032)	-1.270 (1.838)
Openness (1-5)	0.848 (2.713)	-1.823 (1.283)	-0.178 (1.315)
Enjoy Competition (1-5)	-0.708 (1.346)	1.482 (0.675)**	-0.296 (1.372)
Grit (1-5)	-2.178 (5.080)	-3.653 (1.480)**	2.730 (2.189)
Belief on rel. performance (0-1)	26.595 (8.052)***	19.951 (3.870)***	18.982 (3.549)***
Patience Index	-2.860 (2.206)	0.011 (0.754)	0.596 (1.239)
Risk Taking Index	-0.533 (2.740)	0.204 (0.778)	-2.371 (1.190)**
Constant	94.784 (24.931)***	75.462 (14.967)***	79.558 (23.471)***
Num.Obs.	235	458	257
R2 Adj.	0.113	0.474	0.475

Note: Table shows OLS regressions of performance on characteristics in part 2 split by treatment and incentive scheme. Standard errors (in parentheses) clustered on the session level. Productivity (resid.) is a residualized measure of performance in part 1. For this, we regress our full set of predictors on performance. We use the residuals of this regression as a measure of productivity that is corrected for the correlation between performance in Part 1 and the remaining set of our predictors.

Table 6: Productivity by incentive scheme – *Endogenous* treatment

### 3.3 Determinants of Sorting across Payments

Figure 3 shows the number of participants in each payment. The light blue bars refer to the *Exogenous* treatment where the assignment was random, yielding practically the same number of observations for each payment. The dark blue bars for the *Endogenous* treatment reveal that sorting is not random, however. The Variable payment is chosen most often (about half of the time), with the other two payments being roughly similarly attractive and accounting for about a quarter of choices each. Sorting is obviously related to productivity in the 5-minute task in part 1 of the experiment, as the dark blue bars in Figure 4 reveal. Subjects who solved more tables in part 1 are most likely to sort into the Tournament payment and least likely to sort into the Fixed payment. The light blue

bars for the *Exogenous* treatment indicate that performance in part 1 is orthogonal to the random assignment to payments in part 2.

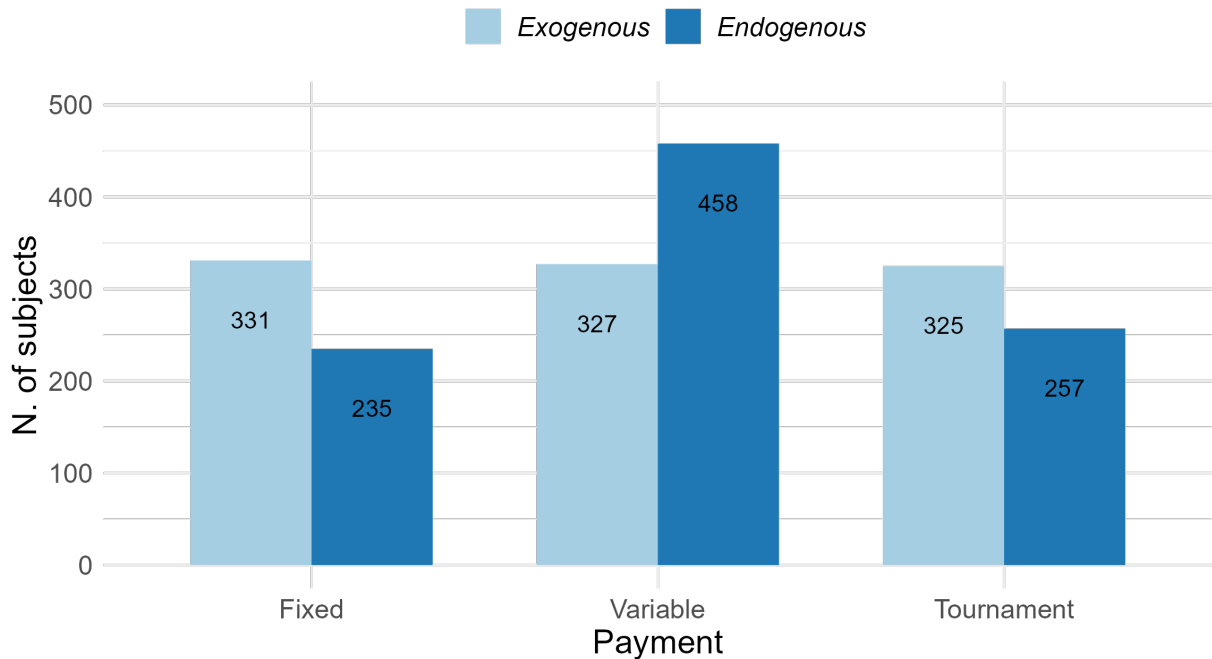


Figure 3: Choices of Payments: Part 2

*Notes:* In the *Endogenous* treatment, a significantly higher number of participants choose Variable over the other two payments (paired  $t$ -test for each difference:  $p < 0.01$ ). The number of participants selecting into the Fixed payment does not differ from the number of participants selecting into the Tournament payment.

Owing to our extensive set of covariates, which is much larger than that of prior studies, we can simultaneously control for all variables (called ‘full model’ below). This comprehensive approach is not feasible in previous studies that concentrate on a limited number of characteristics. Comparing both the ‘full model’ and a simple correlation helps to assess the robustness of our results and shed further light on those reported in the literature.

We provide the resulting overview regarding sorting decision in Table 7 (and report the full logit regressions including coefficients in Table A6). Panel A relates skills to sorting, Panel B demographics, and Panel C personality traits, economic preferences and beliefs. In each panel, the first column lists the respective variables that we consider. The middle column then specifies whether we show our own findings (either in a full model with all variables listed in Table 7 or by only reporting correlations between sorting and the respective variable) or whether we refer to findings in the previous literature. The latter means that Table 7 compares how our findings relate to the ones reported in previous papers that have examined sorting and how it relates to the various variables. None of the papers that we refer to in Table 7 have such a broad range of variables as we have, however. And moreover almost all of them have only pairwise comparisons between two payments (rather than between three as in our case). After the middle column in Table 7

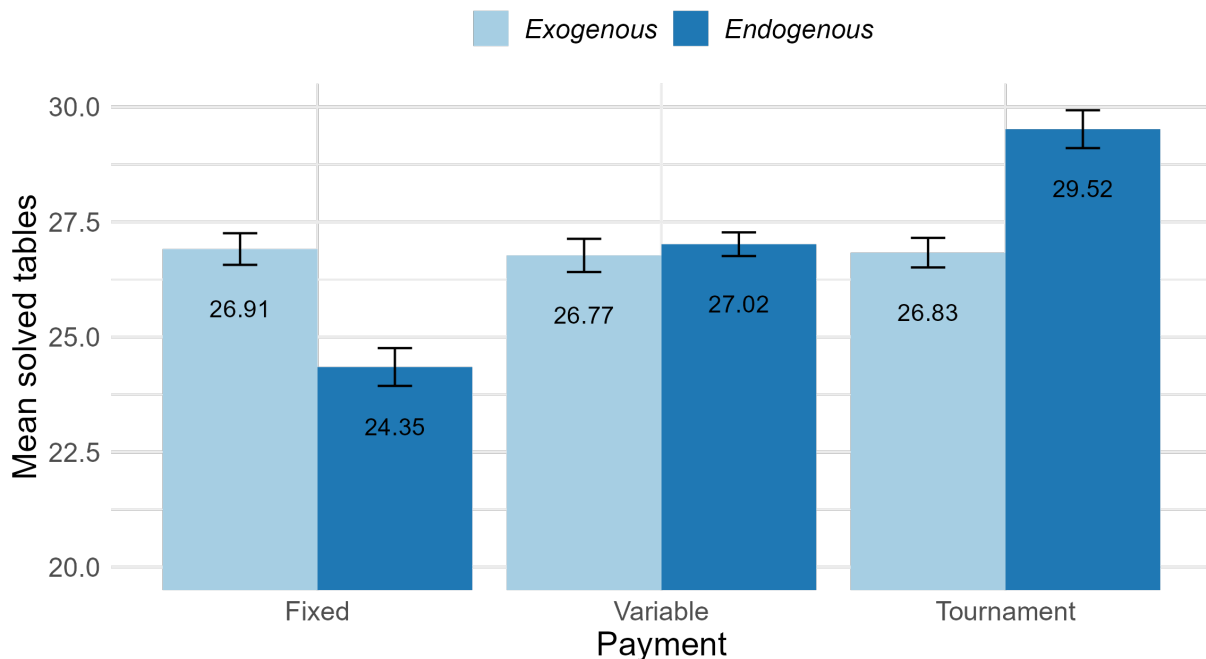


Figure 4: Tasks completed in part 1, by treatment and payment in part 2

*Notes:* In the *Endogenous* treatment, we note that participants choosing the Fixed payment performed the poorest in part 1, followed by the ones choosing the Variable payment, with those choosing the Tournament payment performing best (paired *t*-test for each difference:  $p < 0.01$ ). In the *Exogenous* treatment, there are no such differences in part 1.

we then show the results on the right hand side of the table. The column “Consistent?” indicates whether previous findings are in line with our findings (✓) or not (✗). In the following columns, we show the direction of relationships between sorting and all variables, whether they are significantly positive (▲), significantly negative (▼), or insignificant (○). This is done for all possible pairwise comparisons (with F for the Fixed payment, V for the Variable payment and T for Tournament payment), and in the ultimate column also for an ordered logit model (where the order comes from the riskiness of the outcome, from no risk in the Fixed payment over risk that one has control over in the Variable payment to risk that may even be fully outside one’s own influence in the Tournament payment).

Looking at Panel A, we note first of all that our results are almost always in line with findings of earlier papers (which also holds for Panels B and C, and in the very few other cases, either the relationship in our study or in the reported literature is non significant). The main insight from Panel A is that productivity is essential for sorting (out of the Fixed payment and into the Variable or Tournament payment), which is not surprising.<sup>19</sup> However, the null-findings for IQ came more as a surprise to us, and is the only noticeable deviation from previous literature (Buser et al., 2014). Given that they had only two pay-

<sup>19</sup>Note again that we do not include our baseline measure of productivity from Part 1 as a predictor as it might cover up the potential explanatory power of other predictors. Instead, we use a residualized measure, as explained in Section 2.4.

ments and proxy IQ by the GPA, however, it is not clear what Buser et al. (2014) would have found with three payments as well, and a more direct measure of IQ (also in light of the different sign they report for the math grade). By and large, we also find that better grades in German and math lead to sorting out of the Fixed payment (either in favor of the Variable or the Tournament payment).

Panel B shows that gender is important for sorting, as the large majority of previous papers (albeit with only two payments) has also found. Yet, our results are much more nuanced than what is commonly reported in the literature. Women have been reported to be less likely to sort into competitive payments (Niederle and Vesterlund, 2007; Buser et al., 2014, 2017). Most of the literature has focused on sorting decisions between a variable payment and a tournament payment (Datta Gupta et al., 2005; Niederle and Vesterlund, 2007; Buser et al., 2014; Almås et al., 2016; Buser et al., 2017; Reuben et al., 2017; Buser et al., 2021b,a). Here the relationship is unambiguous. Women shy away from tournament payments more often than men if a variable payment is the alternative. This is also what we find. Dohmen and Falk (2011) find no relation between gender and sorting into a tournament payment over a fixed payment or (separately) between sorting into a variable payment over a fixed payment. With our comparison of three payments, we observe that there is a strong tendency of women to self-select into our Variable payment compared to the Fixed and Tournament payment ( $p < 0.01$ ). This finding is obtained from pairwise correlations as well as partial correlations adjusting for all other predictors, such as risk aversion, and we see this result irrespective of pooling the Fixed payment with the Tournament payment, or only comparing the Variable payment with the Fixed payment. This indicates that by no means women shy away from performance-based payments per se. The other variables captured in Panel B of Table 7 don't show strong relations to sorting. It seems that positive parenting lowers the odds of choosing performance based and increasingly risky payments, and that higher socio-economic status makes participants choose a tournament payment less often, but previous findings – if existing – don't align perfectly with these findings.

Panel C shows results for personality traits, economic preferences and beliefs. Here again our results confirm almost always previous findings, but at the same time extend them by our choice between three different payments (and by the much more encompassing set of explanatory variables). From this part of the table, it becomes clear that the Big Five personality traits matter for sorting (but recall that they hardly mattered for performance in the *Exogenous* treatment). Extraversion and neuroticism are predictive for sorting out of the Variable payment into the Fixed payment ( $p < 0.05$ ). Competitiveness is also an important predictor for sorting, as a higher score in the Competitive Orientation Measure (Newby and Klein, 2014) is related to a higher likelihood to choose the Tourna-

ment payment, as well as for avoiding the Fixed payment ( $p < 0.01$ ).

Beliefs on own relative performance also matter for sorting. Individuals who perceive their own productivity to be on the upper end of the distribution are more likely to sort into the Tournament payment and less likely to choose the Fixed or Variable payment. Also this finding is consistent with prior findings in, e.g., Dohmen et al. (2011). Similarly, we also see – as practically all previous literature – that more risk taking individuals are more likely to sort into the Tournament payment, while they do not seem to matter for the preference between the Variable and Fixed payment.

Both grit and altruism have not been studied in the literature so far with respect to their influence of sorting between payments. For grit we see no relation at all, while more altruism goes hand in hand with a stronger dislike of the Tournament payment.

2

Table 7: Predictors of sorting decisions

Variables	Literature	Consistent?	F $\succ$ V	V $\succ$ F	V $\succ$ T	T $\succ$ V	T $\succ$ F	Ordered logit
<b>Panel A: Skills</b>								
Grade German	<b>Our findings - Full model</b>		✓	○	∧	○	○	∧
	<b>Our findings - Correlation</b>		✓	∧	∧	○	○	∧
Grade Math	<b>Our findings - Full model</b>		○	○	○	∧	∧	∧
	<b>Our findings - Correlation</b>		✓	○	○	∧	∧	∧
IQ	Buser et al. (2014)	✓				∧		
	<b>Our findings - Full model</b>		○	○	○	○	○	○
Productivity (resid.)	<b>Our findings - Correlation</b>		○	○	○	○	○	
	Buser et al. (2014)†	(✗)				∨		
	<b>Our findings - Full model</b>		✓	○	∧	∧	∧	∧
	<b>Our findings - Correlation</b>		✓	○	∧	∧	∧	
	Almås et al. (2016)	✓				∧		
	Dohmen and Falk (2011)§	✓	∨		∧		∧	
	Fornwagner et al. (2023)	✓				∧		
	Niederle and Vesterlund (2007)	✓				∧		
	Reuben et al. (2017)	(✗)				○		
	Sutter and Glätzle-Rützler (2015)	✓				∧		
	Tungodden and Willén (2023)	✓				∧		
	<b>Panel B: Demographics</b>							
Age rel. to grade mean	<b>Our findings - Full model</b>		○	○	○	○	○	○
	<b>Our findings - Correlation</b>		○	○	○	○	○	
Female	<b>Our findings - Full model</b>		∨	∧	∧	∨	∨	○
	<b>Our findings - Correlation</b>		○	∧	∧	∨	∨	○
	Almås et al. (2016)	✓				∨		
	Boneva et al. (2022)	✓				∨		
	Buser et al. (2014)	✓				∨		
	Buser et al. (2017)	✓				∨		
	Buser et al. (2021b)¶	✓				∨		
	Buser et al. (2021a)	✓				∨		
	Buser et al. (2022)	✓				∨		

To be continued on next page ...

Variables	Our Findings/Literature	Consistent?	F $\wedge$ T	V $\wedge$ F or T	V $\wedge$ F	T $\wedge$ V or F	T $\wedge$ V	T $\wedge$ F	Ordered logit	
Grade	Datta Gupta et al. (2005)	✓					✓			
	Dohmen and Falk (2011)§	(X)	○		○			○		
	Eriksson et al. (2009)	(X)					○			
	Fornwagner et al. (2023)	✓					✓			
	Niederle and Vesterlund (2007)	✓					✓			
	Reuben et al. (2017)	✓					✓			
	Sutter and Glätzle-Rützler (2015)	✓					✓			
	Tungodden and Willén (2023)	✓					✓			
	<b>Our findings - Full model</b>			○	○	○	○	○	○	○
	<b>Our findings - Correlation</b>			○	✓	○	○	∧	○	
Number of siblings	Sutter and Glätzle-Rützler (2015)	✓					○			
	<b>Our findings - Full model</b>		○	○	○	○	○	○	○	
Positive parenting	<b>Our findings - Correlation</b>		○	○	○	○	○	○	○	
	<b>Our findings - Full model</b>		○	○	○	○	○	○	○	
Socio-economic status	<b>Our findings - Correlation</b>		○	○	○	○	○	○	○	
	<b>Our findings - Full model</b>		○	○	○	○	○	○	○	
	Almás et al. (2016)	(X)					∧			
	Boneva et al. (2022)§	✓					○			
	Tungodden and Willén (2023)	(X)					✓			
<b>Panel C: Personality Traits, Economic Preferences and Beliefs</b>										
Altruism	<b>Our findings - Full model</b>		○	○	○	○	○	○	○	
	<b>Our findings - Correlation</b>		○	∧	○	✓	✓	✓		
Big 5	<b>Our findings - Full model</b>									
	Extraversion		∧	○	✓	○	○	○	○	
	Agreeableness		○	○	○	○	○	○	○	
	Conscientiousness		○	○	○	○	○	○	○	
	Neuroticism		∧	○	✓	○	○	○	○	
	Openness		○	○	○	○	○	○	○	
	<b>Our findings - Correlation</b>									
	Extraversion		○	✓	○	○	∧	○		
	Agreeableness		○	∧	○	✓	✓	✓		
	Conscientiousness		○	∧	○	○	○	○		
	Neuroticism		∧	○	○	✓	✓	✓		
	Openness		○	○	○	○	○	○		
	Almás et al. (2016)	✓						○		
Extraversion							○			
Agreeableness							○			
Conscientiousness							○			
Neuroticism							○			
Openness							○			
Enjoy Competition	<b>Our findings - Full model</b>		✓	○	∧	∧	∧	∧	∧	
	<b>Our findings - Correlation</b>		✓	✓	○	∧	∧	∧		
Buser et al. (2017)	✓					∧				
Niederle and Vesterlund (2007)	✓					∧				
Reuben et al. (2017)	✓					∧				
Grit	<b>Our findings - Full model</b>		○	○	○	○	○	○	○	
	<b>Our findings - Correlation</b>		○	○	○	○	○	○		
Belief on rel. performance	<b>Our findings - Full model</b>		✓	✓	○	∧	∧	∧	∧	
	<b>Our findings - Correlation</b>		✓	✓	○	∧	∧	∧		
	Almás et al. (2016)	✓					∧			
	Buser et al. (2014)	✓					∧			

To be continued on next page ...

Variables	Our Findings/Literature	Consistent?	F or T	F or T	F	V or F	V	F	Ordered logit	
		✓	∨	∨	∨	∧	∧	∧	∧	
Patience	Buser et al. (2022)¶	✓						∧		
	Datta Gupta et al. (2005)	(X)						○		
	Dohmen and Falk (2011)§	✓	∨			○			∧	
	Fornwagner et al. (2023)	✓						∧		
	Niederle and Vesterlund (2007)	✓						∧		
	Reuben et al. (2017)	✓						∧		
	Sutter and Glätzle-Rützler (2015)	✓						∧		
	Tungodden and Willén (2023)	✓						∧		
	<b>Our findings - Full model</b>			∨	○	○	○	○	○	○
	<b>Our findings - Correlation</b>			∨	○	○	○	○	○	∧
Risk taking	Almås et al. (2016)	○						∧		
	<b>Our findings - Full model</b>		∨	○	○	∧	∧	∧	∧	
	<b>Our findings - Correlation</b>		∨	∨	○	∧	∧	∧		
	Almås et al. (2016)	✓						∧		
	Buser et al. (2014)	✓						∧		
	Buser et al. (2021b)¶	✓						∧		
	Bonin et al. (2007)‡	✓						∧		
	Datta Gupta et al. (2005)	✓						∧		
	Dohmen and Falk (2011)§	✓	∨				∧		∧	
	Eriksson et al. (2009)	✓						∧		
Fornwagner et al. (2023)	(X)						○			
Niederle and Vesterlund (2007)	✓						∧			
Sutter and Glätzle-Rützler (2015)	✓						∧			
Tungodden and Willén (2023)	✓						∧			

**Note:** Table contrasts results from Table A6 with results from literature that demonstrate the role of explanatory skill variables, demographics, personality traits and economic preferences on sorting decisions between the variable (V), fixed (F) and tournament payments (T).  $\wedge$  - Significant increase in likelihood.  $\vee$  - Significant decrease in likelihood.  $\circ$  - No significant results. † - IQ proxied by GPA. Our findings are presented for both, the resp. coefficient for the full model on sorting decisions controlling for all other predictors and the plain correlation between predictor and sorting decision. ¶ - Investigates sorting decisions into college major choices and future earnings. § - Compared sorting decisions between fixed, variable, revenue sharing and tournament payments. § - Compared sorting decisions for a linear and a convex payout schedule that paid more per correct answer. ‡ - Investigates sorting decisions into occupations with low earnings risk.

### 3.4 Identifying the Treatment Effect of Having a Choice

As a final part of this section, we isolate the overall treatment effect of having agency over the payment on performance, and start by contrasting individuals' performance across the *Exogenous* and *Endogenous* treatment conditions. Looking at Figure 2, and recalling the roughly equal choice frequencies of the Fixed and the Tournament payments (see Figure 3), we note that the average performance across all payments is practically the same in *Exogenous* and *Endogenous*. The overall performance is 0.78 units lower, and insignificantly so, in *Endogenous* than in *Exogenous* (121.8 vs. 122.74;  $p = 0.37$ ). This



means that allowing for self-selection into payments is, on average, not beneficial for overall productivity. Of course, Figure 2 reveals that the overall null-effect does not apply to each payment.

Yet of course, the samples under any of the three payments in the *Endogeneous* treatment condition are not comparable to those in the respective payment in the *Exogeneous* treatment condition, and the difference observed between the treatments in Figure 2 with respect to performance is a mixture of having agency over the payment, and different sample characteristics resulting from self selection (starting with productivity, as Figure 4 illustrates).

Causal random forests (Wager and Athey, 2018) allow for a further examination. This type of machine learning algorithm does not only predict outcomes, but, taking into account the causal setting of experiments, predicts outcomes on the level of subgroups based on baseline characteristics – for the scenario with and without having received the treatment, i.e., with and without having the choice over payments. Importantly, individual characteristics and the payment is kept constant. Subgroups are defined in a data-driven way based on explanatory variables and corresponding cut-off values with the goal to maximize between-group heterogeneity of treatment effects. Aggregating the subgroup-based treatment effect estimations offers another way of obtaining an estimate for the conditional average treatment effects for the whole sample, or for certain subgroups, such as those that are linked to the different payments considered here. Effectively, this allows us to compare similar participants in terms of individual characteristics who had self-selected themselves into a given payment with those who had not, but instead have been randomly assigned.

Figure 5 reports these conditional average treatment effects (CATE) conditional on the payments. We find a negative CATE of self-selecting into the Fixed payment: individuals who had agency over the payment and chose the Fixed payment, had a profoundly lower performance than what they are predicted to have based on the performance of individuals with similar characteristics who were randomly assigned to the Fixed payment. For the Variable and Tournament payments we find slightly negative but statistically insignificant differences between performance when self-selected and performance when exogenously assigned to these payments.

- Woerner, A., Romagnoli, G., Probst, B. M., Bartmann, N., Cloughesy, J. N., & Lindemans, J. W. (2021). Should Individuals Choose Their Own Incentives? Evidence from a Mindfulness Meditation Intervention. (R&R ManSci) ([Link](#))
- – Giving subjects choice over incentive schemes backfires. Subjects perform less when being given the choice compared to subjects in the same incentive scheme that were randomly assigned.

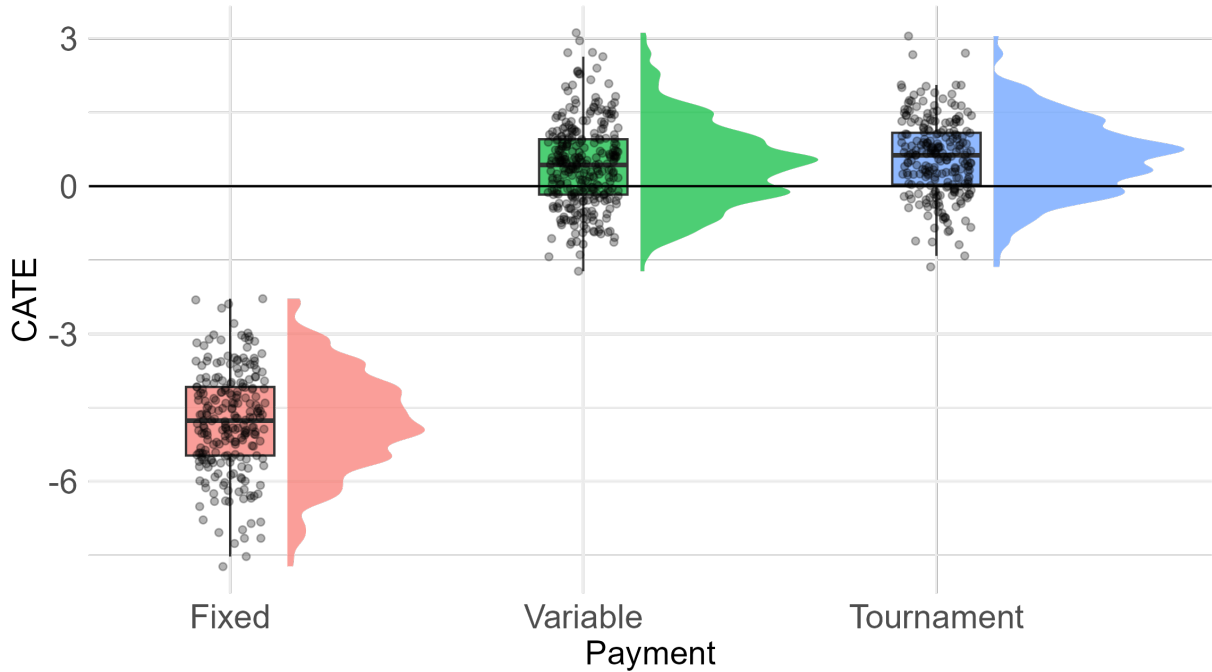


Figure 5: Conditional Average Treatment Effects (CATEs) of Having Agency over the Choice of a Payment on Performance (by Payment).

*Notes:* This figure shows the distribution of CATEs within each payment. Depicted boxplots show 25th/50th/75th percentiles and whiskers for  $1.5 \times \text{IQR}$ . We calculate the CATE by comparing the actual and counterfactual performance in the same payment. Note that the counterfactual case is a prediction based on the causal random forest methodology that predicts individual's performance if the individual had been in the other treatment group and, thus, had (in *Exogenous*) or had not (in *Endogenous*) agency over payments. A CATE of zero indicates no treatment effect of having agency over payments on performance. A negative CATE indicates a negative treatment effect of having agency over payments on performance.

- Dizon-Ross, R., & Zucker, A. (2021). Can price discrimination increase behavioral change? Evidence from a randomized field experiment. Working paper, University of Chicago (Link)
  - Choice over incentive schemes has positive impact on health-behaviors (walking).
- Adjerid, I., Loewenstein, G., Purta, R., & Striegel, A. (2022). Gain-loss incentives and physical activity: the role of choice and wearable health tools. *Management Science*, 68(4), 2642-2667 (Link)
  - Subjects get fitbit wearables and partly have the choice to self-select into incentive schemes aiming for better health-behaviors. Choice backfires for some: subjects in the same incentive schemes w/o the choice achieve better health outcomes.

Finally, we further utilize the preceding analysis using causal random forests and seek to identify the factors that let participants choose a payment that is not performance-maximizing. By doing so, we tackle the question for whom agency on own incentivization

is harmful from a performance-maximizing perspective. For this, we analyze the heterogeneity using a best linear projection on the previously predicted individual-level CATE in Table 8. Controlling for payments, we find that participants who are relatively older or are more productive in Part 1 are predicted to profit from agency in own incentivization in line with a performance-maximizing objective. In contrast, participants who are extrovert or impatient tend to sort inefficiently.

	CATE (pred.)
	(1)
<b>Skills</b>	
Productivity Part 1	0.612*** (0.226)
<b>Demographics</b>	
Age rel. to grade mean	4.494*** (1.546)
Grade (9-13)	-3.247 (2.182)
Positive Parenting (1-5)	0.657 (1.210)
Extraversion (1-5)	-3.470** (1.465)
<b>Personality Traits, Economic Preferences, and Beliefs</b>	
Agreeableness (1-5)	2.752 (2.057)
Belief on rel. performance (0-1)	-7.089 (5.045)
Patience Index	-3.101*** (1.164)
Risk Taking Index	-0.066 (1.897)
Variable	11.434*** (3.038)
Tournament	12.312*** (3.403)
Constant	15.325 (27.437)
<b>Num.Obs.</b>	1160

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

*Note:*

Best linear projection of the conditional average treatment effect. Confidence intervals are cluster- and heteroskedasticity-robust. Predictors used represent the top 10 most important variables for the random forest prediction.

Table 8: Heterogeneity in CATEs on Performance – Best linear Projection

## 4 The Potential of Algorithmic Assignment to Payments

Here we investigate whether and to which extent the exogenous assignment to payments could be improved, given the knowledge about which factors (with respect to personal traits, preferences, and skills) determine productivity. Recall that Opitz et al. (forthcoming) first ran an exogenous treatment – like ours – and then used the insights from this treatment for an endogenous assignment of a new set of MTurkers to the payment that was predicted to yield the best output. In fact, they found that algorithmic assignment improves performance beyond the best exogenous treatment. Our approach is different, but complementary. We use machine learning algorithms to estimate whether we could have improved performance, payoffs, and utility through such an assignment. So, we do not run additional sessions with new treatments, but we consider three different outcomes, and in particular include outcomes that focus on the employee’s perspective, which might be an important aspect to consider in light of potential backlash to the use of algorithmic tools in human resource practice. For a measure of utility, we assume a standard utility function being defined as the difference between absolute payoff and effort costs,  $U(e) = \pi(e) - c(e)$  (see, e.g., DellaVigna and Pope, 2017; DellaVigna et al., 2022). To elicit effort costs  $c(e)$ , we have asked all individuals how much effort they exerted, how stressed they felt, and how exhausted they got (all 1-7 Likert scales). These questions were asked right after the 5-minute real effort ask in Part 1. We define effort costs  $c(e)$  as the mean of all three responses on the individual level standardized by the sample standard deviation across all three responses.

We run random forests trained on individuals in the *Exogenous* treatment to identify – without any sorting – influential predictors of our outcomes, i.e., performance, payoff, and utility, in each payment. In a second step, we use these insights to predict each outcome for each individual in the *Endogenous* condition under the two unobserved payments. This serves to get a prediction of each unobserved outcome if participants in *Endogenous* had been assigned exogenously to another payment. Lastly, by comparing realized outcomes in the actually chosen payment with our random forest predictions for a given set of characteristics we isolate one “optimal” payment where each outcome is maximized on the individual level. The resulting differences between the realized outcomes in the actually chosen and the predicted outcome in the optimal payment – if the two do not coincide – allow us to quantify the discrepancies in outcomes due to possibly suboptimal sorting decisions. If the chosen and optimal payment are identical, the difference is zero.

Figure 6 provides an overview of algorithmic sorting decisions resulting from this random forest based approach. In the left panel, we show the results for performance, in

the middle one for payoffs, and in the right one for utility. Within each panel, the first column refers to participants that have chosen the Fixed payment, the second to those in the Variable payment, and the third column to participants in the Tournament payment. In each column, we report which would have been the optimal choice (for each of the three outcomes). In the first panel and the first column, we see, for example, that for only 20% of participants who have chosen the Fixed payment, this was estimated to be the optimal choice for maximizing performance. For 43%, choosing the Variable payment would have led to higher performance, and more than one third (37%) should have chosen the Tournament payment to maximize performance. For the group having chosen the Variable payment (second column), we note that 58% have maximized the estimated performance with this payment (which is much better than the low 20% for participants in the Fixed payment). For the group who chose the Tournament payment, almost two thirds (65%) are estimated to have taken the optimal choice.

The middle panel shows that with regards to potential payoffs, the fraction of optimal choices is smaller than for performance in the first panel. This is due to the fact that the payment with the highest estimated performance of an individual need not be the one with the highest payoffs, because the latter depend on the absolute level of performance (for the Variable payment) and the relative performance compared to other participants (in the Tournament payment). Here we see the largest fraction of optimal choice for the Tournament payment, where 52% of participants are estimated to have actually chosen the payment that maximizes their payoff.

The right panel in Figure 6 shows the estimates for utility. Remarkably, 43% of participants who chose the Fixed payment are estimated to have optimized their utility (while performance was optimal only for 20% and payoffs only for 7% of these participants). This shows that participants in the Fixed payment seem to have perceived the task as relatively costly, thus improving their utility by reducing effort, and consequently performance and payoffs. For subjects who have chosen the Variable or the Tournament payment, we see that they optimized their utility in 42%, respectively 51%, of cases.

Across all three outcome measures and all groups of participants (those choosing Fixed, Variable and Tournament payment), we observe that on average about 57.4% of participants make sorting decisions that fall short of their predicted potential<sup>20</sup>. This is a substantial share of participants, indicating a large potential for improving outcomes by taking into account personal characteristics when assigning subjects to different payments. We can quantify this potential as follows: Figure 7 reports results on the differences between actual and predicted outcomes. Note that we present differences in a standardized form, i.e., we divide differences by their sample standard deviation. We find that the algorithmic assignment of payments is in every combination of payment and outcome able to assign

---

<sup>20</sup>To arrive at 57.4%, we add all shares of participants that did not choose the outcome-maximizing payment across all payment-outcome combinations and divide it nine.

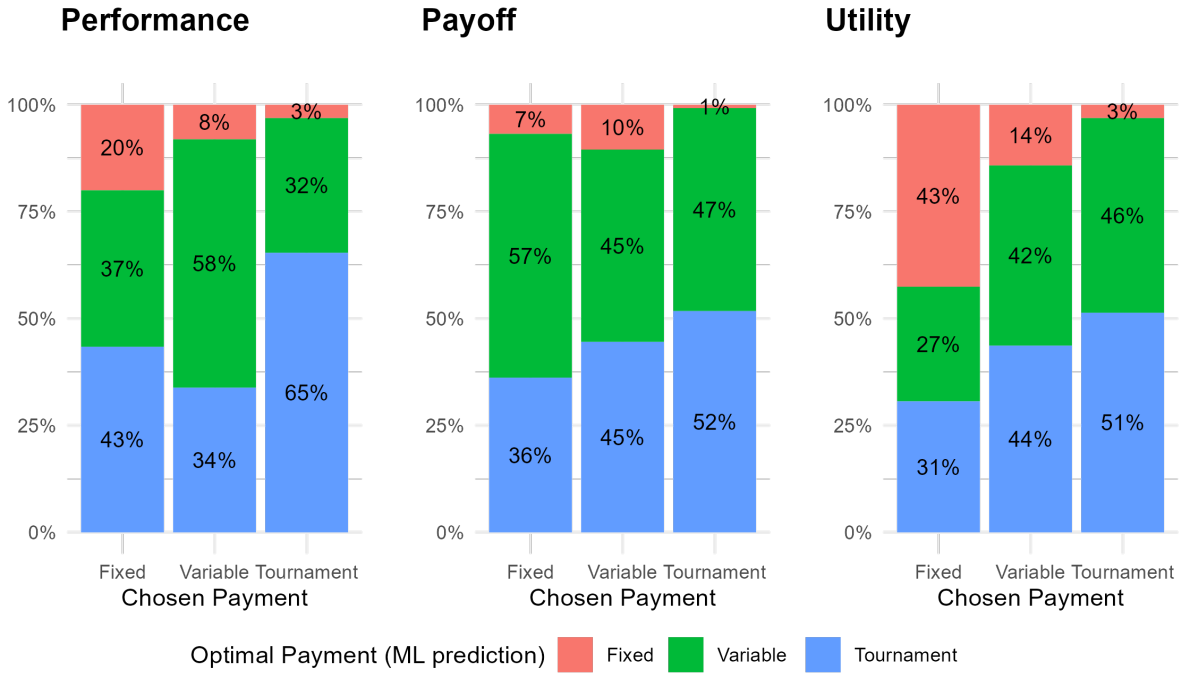


Figure 6: Algorithmic assignment into payoff-, performance- and utility-maximizing payment based on random forest predictions compared to outcomes resulting from own choice

individuals into payments where the average predicted outcome is significantly higher than the actually realized average outcome. This holds for all three outcomes, meaning that the algorithmic approach would not only improve performance and payoffs, but also utility on average, meaning that such an approach would not only be beneficial for potential employers (who clearly prefer better performance), but also for the participants (with higher average payoffs and utility).

## 5 Conclusion

Improving performance through proper incentives sounds like a simple solution to the important question of how to motivate humans to perform at their best. Yet, it is not the case that one size fits all, nor does one payment yield the best performance. Rather, humans react in very different ways to incentives, and this reaction depends on many factors, such as their abilities, background characteristics, personality traits, economic preferences and also their beliefs. For this reason, it is important to understand how these factors interact with different payments for performance, payoffs and utility from a task.

In this paper, we have studied in a lab-in-the-field experiment with 1,914 German high school students how a large plethora of factors influence sorting decisions and performance in a Fixed, a Variable and a Tournament payment. We have found evidence of systematic sorting across the three payments based on socio-demographics, personality traits, preferences and beliefs. Interestingly, the factors that explain sorting are only partly overlapping

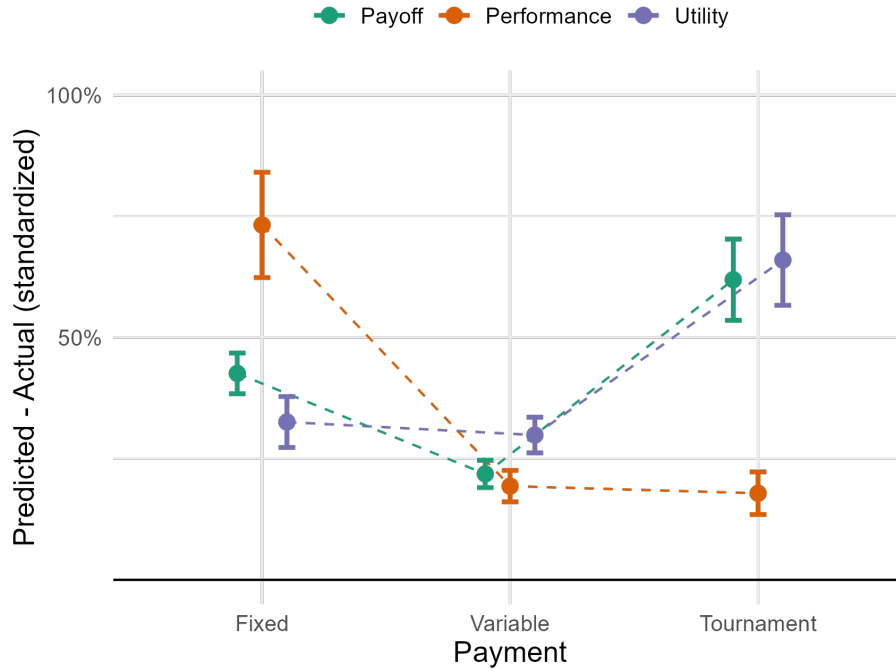


Figure 7: Potential Improvement through Algorithmic Assignment: Standardized difference between predicted and actual outcomes (payoff, performance, and utility) across payments. Predicted values are based on algorithmic assignment into the outcome-maximizing payment. Standardization results from dividing the difference between predicted and actual outcomes by the population standard deviation of this difference. Brackets represent 95%-confidence intervals.

with the factors that have explanatory power for performance. Moreover, the importance of the different factors also depends on which of the payments are compared to each other.

We also show that specific factors determine effort provision, but differently depending on the payment. So, there is a large degree of heterogeneity in the reaction to incentives, which is the reason why one size cannot fit all. When given the choice to self-select into a specific payment, we observe that individuals' choice behavior is often not in line with the factors that would maximize their effort provision, payoffs or utility from the task. In fact, about 50% of sorting decisions could have been improved by a machine learning algorithm that is trained on the relationship between personal characteristics and performance when incentives are exogenously assigned. As abilities, beliefs about own abilities, socio-demographics, personality traits, and preferences are impacting sorting decisions and effort provision in different ways, depending on the payment, it looks promising for future research to investigate these intricate interdependencies further. This may then help in raising awareness on how to improve labor market contracts, where the employer as well as the worker, can better account for the strengths and weaknesses of characteristics to optimize output and earnings, but also utility, under different payments.

## References

- Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman**, “Reference Points and Effort Provision,” *American Economic Review*, 2011, *101* (2), 470–92.
- Akee, Randall, William Copeland, E Jane Costello, and Emilia Simeonova**, “How Does Household Income Affect Child Personality Traits and Behaviors?,” *American Economic Review*, 2018, *108* (3), 775–827.
- Alan, Sule and Seda Ertac**, “Fostering Patience in the Classroom: Results from Randomized Educational Intervention,” *Journal of Political Economy*, 2018, *126* (5), 1865–1911.
- , **Teodora Boneva, and Seda Ertac**, “Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit,” *The Quarterly Journal of Economics*, 2019, *134* (3), 1121–1162.
- Almås, Ingvild, Alexander W Cappelen, Kjell G Salvanes, Erik Ø Sørensen, and Bertil Tungodden**, “Willingness to Compete: Family Matters,” *Management Science*, 2016, *62* (8), 2149–2162.
- Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz**, “Personality Psychology and Economics,” in “Handbook of the Economics of Education,” Vol. 4, Elsevier, 2011, pp. 1–181.
- Andersen, Anette, Rikke Krolner, Candace Currie, Lorenza Dallago, Pernille Due, Matthias Richter, Ágota Örkényi, and Bjorn Holstein**, “High Agreement on Family Affluence Between Children’s and Parents’ Reports: International Study of 11-year olds,” *Journal of Epidemiology & Community Health*, 2008.
- Ariely, Dan, Uri Gneezy, George Loewenstein, and Nina Mazar**, “Large Stakes and Big Mistakes,” *The Review of Economic Studies*, 2009, *76* (2), 451–469.
- Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse**, “The Relationship Between Economic Preferences and Psychological Personality Measures,” *Annual Review of Economics*, 2012, *4* (1), 453–478.
- Boneva, Teodora, Thomas Buser, Armin Falk, and Fabian Kosse**, “The origins of gender differences in competitiveness and earnings expectations: Causal evidence from a mentoring intervention,” *CEPR discussion paper no. DP17008*, 2022.
- Bonin, Holger, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde**, “Cross-sectional Earnings Risk and Occupational Sorting: The Role of Risk Attitudes,” *Labour Economics*, 2007, *14* (6), 926–937.



- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer**, “Beliefs about Gender,” *American Economic Review*, 2019, 109 (3), 739–73.
- Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel**, “The Economics and Psychology of Personality Traits,” *Journal of Human Resources*, 2008, 43 (4), 972–1059.
- , **Huib Meijers, and Bas Ter Weel**, “The Role of Noncognitive Skills in Explaining Cognitive Test Scores,” *Economic Inquiry*, 2008, 46 (1), 2–12.
- Boyce, William, Torbjorn Torsheim, Candace Currie, and Alessio Zambon**, “The Family Affluence Scale as a Measure of National Wealth: Validation of an Adolescent Selfreport Measure,” *Social Indicators Research*, 2006, 78 (3), 473–487.
- Buser, Thomas, Alexander W Cappelen, and Bertil Tungodden**, “Fairness and Willingness to Compete,” *NHH Dept. of Economics Discussion Paper*, 2021, (08).
- , **Muriel Niederle, and Hessel Oosterbeek**, “Gender, Competitiveness, and Career Choices,” *The Quarterly Journal of Economics*, 2014, 129 (3), 1409–1447.
- , – , **and –** , “Can competitiveness predict education and labor market outcomes? Evidence from incentivized choice and survey measures,” Technical Report, National Bureau of Economic Research 2021.
- , **Noemi Peter, and Stefan C Wolter**, “Gender, Competitiveness, and Study Choices in High School: Evidence from Switzerland,” *American Economic Review*, 2017, 107 (5), 125–30.
- , – , **and –** , “Willingness to compete, gender and career choices along the whole ability distribution,” *Experimental economics*, 2022, 25 (5), 1299–1326.
- Cadena, Brian C and Benjamin J Keys**, “Human Capital and the Lifetime Costs of Impatience,” *American Economic Journal: Economic Policy*, 2015, 7 (3), 126–53.
- Carneiro, Pedro Manuel and James J Heckman**, “Human Capital Policy,” *IZA Discussion Paper*, 2003, (821).
- Cassar, Lea and Stephan Meier**, “Nonmonetary Incentives and the Implications of Work as a Source of Meaning,” *Journal of Economic Perspectives*, 2018, 32 (3), 215–38.
- Cawley, John, James Heckman, and Edward Vytlačil**, “Three Observations on Wages and Measured Cognitive Ability,” *Labour Economics*, 2001, 8 (4), 419–442.
- Charness, Gary, Uri Gneezy, and Austin Henderson**, “Experimental Methods: Measuring Effort in Economics Experiments,” *Journal of Economic Behavior & Organization*, 2018, 149, 74–87.

- Chen, Daniel L, Martin Schonger, and Chris Wickens**, “oTree – an Open-source Platform for Laboratory, Online, and Field Experiments,” *Journal of Behavioral and Experimental Finance*, 2016, 9, 88–97.
- Cubel, Maria, Ana Nuevo-Chiquero, Santiago Sanchez-Pages, and Marian Vidal-Fernandez**, “Do Personality Traits Affect Productivity? Evidence from the Laboratory,” *The Economic Journal*, 2016, 126 (592), 654–681.
- Cunha, Flavio and James Heckman**, “The Technology of Skill Formation,” *American Economic Review*, 2007, 97 (2), 31–47.
- Deckers, Thomas, Armin Falk, Fabian Kosse, Pia R Pinger, and Hannah Schildberg-Hörisch**, “Socio-economic status and inequalities in children’s IQ and economic preferences,” *IZA discussion paper*, 2017.
- DellaVigna, Stefano and Devin Pope**, “What Motivates Effort? Evidence and Expert Forecasts,” *The Review of Economic Studies*, 2017, 85 (2), 1029–1069.
- **and M Daniele Paserman**, “Job Search and Impatience,” *Journal of Labor Economics*, 2005, 23 (3), 527–588.
- , **John A List, Ulrike Malmendier, and Gautam Rao**, “Estimating social preferences and gift exchange at work,” *American Economic Review*, 2022, 112 (3), 1038–1074.
- Deming, David J**, “The Growing Importance of Social Skills in the Labor Market,” *The Quarterly Journal of Economics*, 2017, 132 (4), 1593–1640.
- Dobbie, Will and Roland G Fryer Jr**, “The Medium-term Impacts of High-achieving Charter Schools,” *Journal of Political Economy*, 2015, 123 (5), 985–1037.
- Doepke, Matthias, Giuseppe Sorrenti, and Fabrizio Zilibotti**, “The Economics of Parenting,” *Annual Review of Economics*, 2019, 11.
- Dohmen, Thomas and Armin Falk**, “You Get What You Pay For: Incentives and Selection in the Education System,” *The Economic Journal*, 2010, 120 (546), F256–F271.
- **and –**, “Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender,” *American Economic Review*, 2011, 101 (2), 556–90.
- , – , **David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner**, “Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences,” *Journal of the European Economic Association*, 2011, 9 (3), 522–550.
- Dohmen, Thomas J**, “Do professionals choke under pressure?,” *Journal of Economic Behavior & Organization*, 2008, 65 (3-4), 636–653.

- Donato, Katherine, Grant Miller, Manoj Mohanan, Yulya Truskinovsky, and Marcos Vera-Hernández**, “Personality Traits and Performance Contracts: Evidence from a Field Experiment among Maternity Care Providers in India,” *American Economic Review*, 2017, 107 (5), 506–10.
- Duckworth, Angela L, Christopher Peterson, Michael D Matthews, and Dennis R Kelly**, “Grit: Perseverance and Passion for Long-term Goals,” *Journal of Personality and Social Psychology*, 2007, 92 (6), 1087.
- Duckworth, Angela Lee and Patrick D Quinn**, “Development and Validation of the Short Grit Scale (GRIT-S),” *Journal of Personality Assessment*, 2009, 91 (2), 166–174.
- Eriksson, Tor, Sabrina Teyssier, and Marie-Claire Villeval**, “Self-selection and the Efficiency of Tournaments,” *Economic Inquiry*, 2009, 47 (3), 530–548.
- Essau, Cecilia A, Satoko Sasagawa, and Paul J Frick**, “Psychometric Properties of the Alabama Parenting Questionnaire,” *Journal of Child and Family Studies*, 2006, 15 (5), 595–614.
- Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde**, “Global Evidence on Economic Preferences,” *The Quarterly Journal of Economics*, 2018, 133 (4), 1645–1692.
- , – , – , **David Huffman, and Uwe Sunde**, “The preference survey module: A validated instrument for measuring risk, time, and social preferences,” *Management Science*, 2023, 69 (4), 1935–1950.
- Fornwagner, Helena, Monika Pompeo, and Nina Serdarevic**, “Choosing competition on behalf of someone else,” *Management Science*, 2023, 69 (3), 1555–1574.
- Frick, Paul J**, “The Alabama Parenting Questionnaire,” *University of Alabama*, 1991.
- Fuchs, Thomas and Ludger Woessmann**, “What Accounts for International Differences in Student Performance? A Re-examination Using PISA Data,” in “The Economics of Education and Training,” Springer, 2008, pp. 209–240.
- Gaudecker, Hans-Martin Von, Arthur Van Soest, and Erik Wengstrom**, “Heterogeneity in Risky Choice Behavior in a Broad Population,” *American Economic Review*, 2011, 101 (2), 664–94.
- Gneezy, Uri and Pedro Rey-Biel**, “On the Relative Efficiency of Performance Pay and Noncontingent Incentives,” *Journal of the European Economic Association*, 2014, 12 (1), 62–72.

- , **John A List, Jeffrey A Livingston, Xiangdong Qin, Sally Sadoff, and Yang Xu**, “Measuring Success in Education: The Role of Effort on the Test Itself,” *American Economic Review: Insights*, 2019, 1 (3), 291–308.
- , **Muriel Niederle, and Aldo Rustichini**, “Performance in Competitive Environments: Gender Differences,” *The Quarterly Journal of Economics*, 2003, 118 (3), 1049–1074.
- , **Stephan Meier, and Pedro Rey-Biel**, “When and Why Incentives (Don’t) Work to Modify Behavior,” *Journal of Economic Perspectives*, 2011, 25 (4), 191–210.
- Golsteyn, Bart HH, Hans Grönqvist, and Lena Lindahl**, “Adolescent time preferences predict lifetime outcomes,” *The Economic Journal*, 2014, 124 (580), F739–F761.
- Gupta, Nabanita Datta, Anders Poulsen, and Marie Claire Villeval**, “Male and Female Competitive Behavior - Experimental Evidence,” *GATE Working Paper*, 2005, (05-12).
- Hansson, Åse and Jan-Eric Gustafsson**, “Measurement Invariance of Socioeconomic Status Across Migrational Background,” *Scandinavian Journal of Educational Research*, 2013, 57 (2), 148–166.
- Hanushek, Eric A and Ludger Woessmann**, “The Economics of International Differences in Educational Achievement,” in “Handbook of the Economics of Education,” Vol. 3, Elsevier, 2011, pp. 89–200.
- , **Stephen J Machin, and Ludger Woessmann**, *Handbook of the Economics of Education*, Elsevier, 2016.
- Hartley, Jane EK, Kate Levin, and Candace Currie**, “A New Version of the HBSC Family Affluence Scale-fAS Iii: Scottish Qualitative Findings from the International FAS Development Study,” *Child Indicators Research*, 2016, 9 (1), 233–245.
- Heckman, James J**, “Skill formation and the economics of investing in disadvantaged children,” *Science*, 2006, 312 (5782), 1900–1902.
- , “The Economics, Technology, and Neuroscience of Human Capability Formation,” *Proceedings of the National Academy of Sciences*, 2007, 104 (33), 13250–13255.
- **and Yona Rubinstein**, “The Importance of Noncognitive Skills: Lessons from the GED Testing Program,” *American Economic Review*, 2001, 91 (2), 145–149.
- John, Oliver P and Sanjay Srivastava**, “The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives,” *Handbook of Personality: Theory and Research*, 1999, 2 (1999), 102–138.

- Kosse, Fabian and Michela M Tincani**, “Prosociality predicts labor market success around the world,” *Nature communications*, 2020, *11* (1), 5298.
- , **Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk**, “The Formation of Prosociality: Causal Evidence on the Role of Social Environment,” *Journal of Political Economy*, 2020, *128* (2), 434–467.
- Larkin, Ian and Stephen Leider**, “Incentive Schemes, Sorting, and Behavioral Biases of Employees: Experimental Evidence,” *American Economic Journal: Microeconomics*, 2012, *4* (2), 184–214.
- Lindqvist, Erik and Roine Vestman**, “The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment,” *American Economic Journal: Applied Economics*, 2011, *3* (1), 101–28.
- Newby, Jennifer L and Rupert G Klein**, “Competitiveness Reconceptualized: Psychometric Development of the Competitiveness Orientation Measure as a Unified Measure of Trait Competitiveness,” *The Psychological Record*, 2014, *64* (4), 879–895.
- Niederle, Muriel and Lise Vesterlund**, “Do Women Shy Away from Competition? Do Men Compete Too Much?,” *The Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.
- and – , “Explaining the Gender Gap in Math Test Scores: The Role of Competition,” *Journal of Economic Perspectives*, 2010, *24* (2), 129–44.
- Opitz, Saskia, Dirk Sliwka, Timo Vogelsang, and Tom Zimmermann**, “The targeted assignment of incentive schemes,” *Management Science*, forthcoming.
- Prendergast, Canice**, “The Provision of Incentives in Firms,” *Journal of Economic Literature*, 1999, *37* (1), 7–63.
- Raven, John**, “The Raven’s Progressive Matrices: Change and Stability over Culture and Time,” *Cognitive Psychology*, 2000, *41* (1), 1–48.
- Reuben, Ernesto, Matthew Wiswall, and Basit Zafar**, “Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender,” *The Economic Journal*, 2017, *127* (604), 2153–2186.
- Riener, Gerhard, Sebastian O Schneider, and Valentin Wagner**, “Addressing validity and generalizability concerns in field experiments,” *MPI Collective Goods Discussion Paper*, 2020, (2020/16).
- Rutkowski, David and Leslie Rutkowski**, “Measuring Socioeconomic Background in PISA: One Size Might Not Fit All,” *Research in Comparative and International Education*, 2013, *8* (3), 259–278.

- Schneider, Sebastian O and Martin Schlather**, “A New Approach to Treatment Assignment for One and Multiple Treatment Groups,” *CRC Discussion Papers*, 2017, (228).
- Schulz, Wolfram**, “Measuring the Socio-economic Background of Students and Its Effect on Achievement on Pisa 2000 and Pisa 2003,” *Annual Meetings of the American Educational Research Association*, 2006.
- Segal, Carmit**, “Working When No One Is Watching: Motivation, Test Scores, and Economic Success,” *Management Science*, 2012, *58* (8), 1438–1457.
- Sutter, Matthias and Daniela Glätzle-Rützler**, “Gender differences in the willingness to compete emerge early in life and persist,” *Management Science*, 2015, *61* (10), 2339–2354.
- , **Martin G. Kocher, Daniela Glätzle-Rützler, and Stefan T. Trautmann**, “Impatience and Uncertainty: Experimental Decisions Predict Adolescents’ Field Behavior,” *American Economic Review*, 2013, *103* (1), 510–31.
- Torsheim, Torbjørn, Franco Cavallo, Kate Ann Levin, Christina Schnohr, Joanna Mazur, Birgit Niclasen, and Candace Currie**, “Psychometric Validation of the Revised Family Affluence Scale: A Latent Variable Approach,” *Child Indicators Research*, 2016, *9* (3), 771–784.
- Tungodden, Jonas and Alexander Willén**, “When parents decide: Gender differences in competitiveness,” *Journal of Political Economy*, 2023, *131* (3), 751–801.
- Wager, Stefan and Susan Athey**, “Estimation and inference of heterogeneous treatment effects using random forests,” *Journal of the American Statistical Association*, 2018, *113* (523), 1228–1242.
- West, Martin R and Ludger Woessmann**, “‘Every Catholic Child in a Catholic School’: Historical Resistance to State Schooling, Contemporary Private Competition and Student Achievement Across Countries,” *The Economic Journal*, 2010, *120* (546), F229–F255.
- Woessmann, Ludger**, “The Importance of School Systems: Evidence from International Differences in Student Achievement,” *Journal of Economic Perspectives*, 2016, *30* (3), 3–32.

# A Appendix

## A.1 Map of participating schools

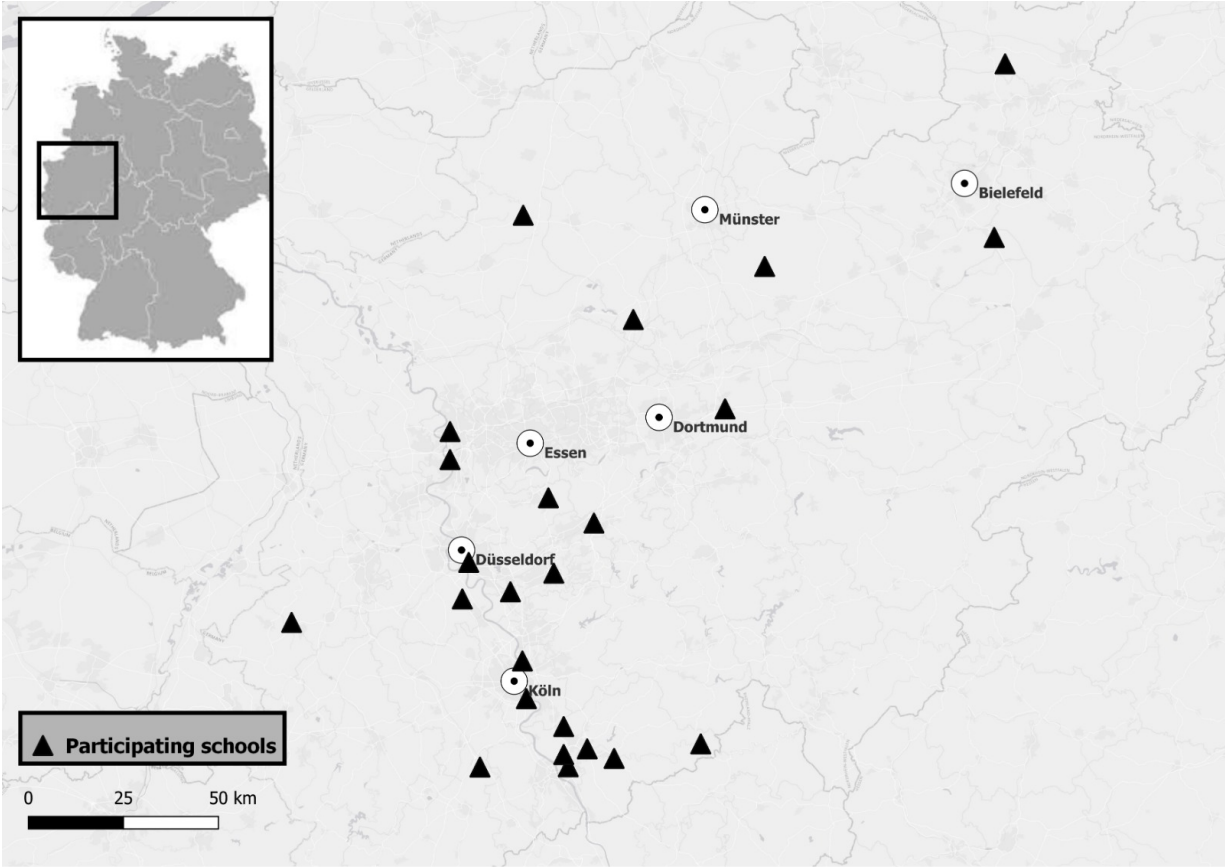


Figure A1: Map of participating schools

## A.2 Additional Results

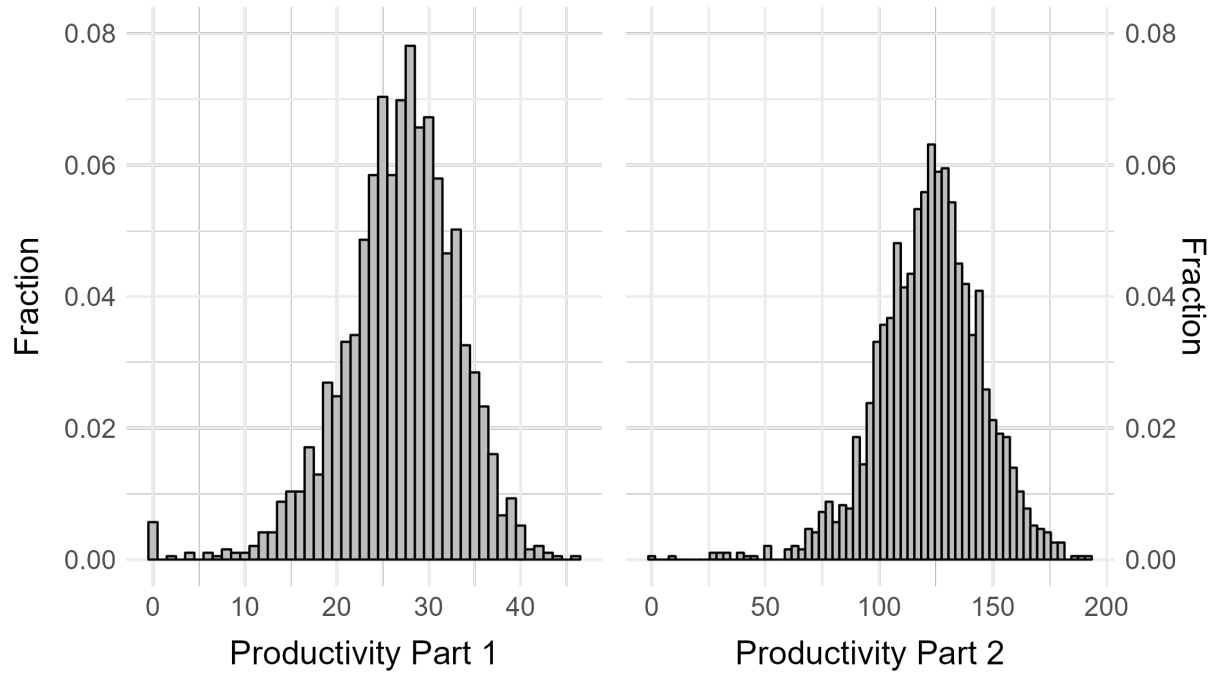


Figure A2: Histogram of solved tables in part 1 and 2.



Characteristic	Overall	Full Observation	Dropout	p-value
<b>Skills</b>				
Grade German	2.73 (0.92)	2.73 (0.91)	2.76 (0.97)	0.39
Grade Math	2.83 (1.17)	2.81 (1.16)	3.01 (1.23)	0.01
IQ (Raven 0-10)	5.05 (1.50)	5.08 (1.48)	4.89 (1.63)	0.13
Productivity part 1 (RET 5 min)	26.80 (6.44)	26.93 (6.24)	25.94 (7.55)	0.15
<b>Demographics</b>				
Age rel. to grade mean	0.00 (0.82)	-0.01 (0.81)	0.08 (0.91)	0.05
Female (=1)	0.53 (0.50)	0.53 (0.50)	0.53 (0.50)	0.93
Grade (9-13)	11.29 (1.00)	11.29 (1.00)	11.29 (1.00)	0.96
Number of siblings	1.64 (1.14)	1.64 (1.12)	1.64 (1.25)	1.00
Positive parenting (1-5)	3.39 (0.92)	3.39 (0.91)	3.41 (0.96)	0.74
FAS Index	6.63 (2.09)	6.68 (1.97)	6.28 (2.73)	0.20
PISA Index	12.76 (2.72)	12.83 (2.37)	12.29 (4.30)	0.98
Family Education Index	0.62 (0.49)	0.61 (0.49)	0.65 (0.48)	0.20
<b>Personality Traits and Economic Preferences</b>				
Altruism (0-10)	7.42 (2.32)	7.49 (2.22)	6.97 (2.85)	0.03
Extraversion (1-5)	3.48 (0.78)	3.48 (0.78)	3.47 (0.79)	0.90
Agreeableness (1-5)	3.56 (0.57)	3.56 (0.58)	3.50 (0.56)	0.06
Conscientiousness (1-5)	3.34 (0.41)	3.35 (0.40)	3.30 (0.44)	0.09
Neuroticism (1-5)	2.91 (0.72)	2.90 (0.72)	2.99 (0.70)	0.03
Openness (1-5)	3.43 (0.66)	3.43 (0.65)	3.42 (0.69)	0.86
Enjoy competition (1-5)	3.10 (1.00)	3.11 (1.00)	3.05 (1.02)	0.45
Grit (1-5)	3.21 (0.50)	3.23 (0.45)	3.09 (0.74)	0.05
Patience Index	0.00 (0.80)	-0.01 (0.80)	0.05 (0.79)	0.26
Risk Index	0.00 (0.75)	0.01 (0.75)	-0.05 (0.76)	0.21

Note: The p-values report results from Wilcoxon rank-sum tests of differences between the two groups. We can not compare the Composite SES index across fully matched and non-matched observations as the Composite SES index is constructed by means of a principal component analyses using only fully matched observations. Thus, for an analyses of attrition along socio-economic status, we substitute the Composite SES index by the FAS, index, and Family Education indices. For a detailed description of how we construct these indices, please see Section A.4.

Table A1: Attrition between complete observations and dropouts after first session

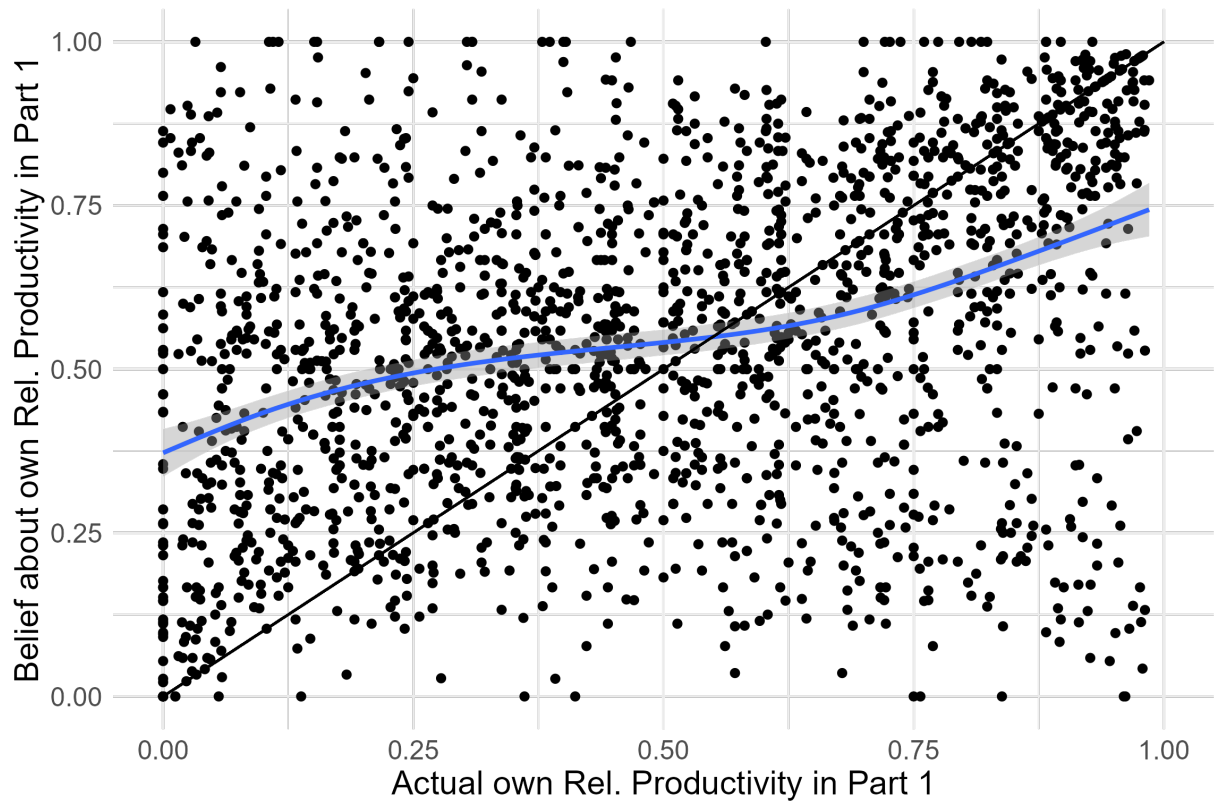


Figure A3: Scatterplot contrasting beliefs on relative productivity and actual productivity in part 1. Values above (below) diagonal black line represent overconfident (underconfident) self-assessments.

### A.3 Regression analysis

	Using Fixed as base
	(1)
Variable	-5.084 (24.515)
Tournament	-11.027 (22.866)
Productivity (resid.)	1.711 (0.284)***
Female (=1)	1.614 (2.826)
Age rel. to grade mean	-0.576 (0.987)
Grade (9-13)	2.315 (1.460)
IQ (Raven 0-10)	0.909 (0.739)
Grade Math	3.232 (0.813)***
Grade German	-1.182 (0.966)
Number of siblings	-0.495 (0.906)
Composite SES Index	0.203 (0.817)
Patience Index	-0.255 (1.263)
Risk Index	1.110 (1.316)
Altruism (0-10)	0.518 (0.576)
Extraversion (1-5)	-2.868 (1.293)**
Agreeableness (1-5)	-1.424 (2.038)
Conscientiousness (1-5)	4.135 (2.731)
Neuroticism (1-5)	-0.256 (1.592)
Openness (1-5)	-0.200 (1.346)
Enjoy Competition (1-5)	0.681 (0.970)
Positive Parenting (1-5)	-1.151 (1.186)
Grit (1-5)	-2.522 (2.127)
Belief on rel. performance (0 - 1)	26.292 (4.537)***
Variable x Productivity (resid.)	0.353 (0.286)
Variable x Female (=1)	1.871 (3.296)
Variable x Rel. age grade mean	-1.202 (1.071)
Variable x Grade (9-13)	0.478 (1.632)
Variable x IQ (Raven 0-10)	-0.219 (0.755)
Variable x Grade Math	-1.063 (0.950)
Variable x Grade German	0.599 (1.265)
Variable x Number of siblings	0.246 (1.078)
Variable x Composite SES Index	-0.783 (0.906)
Variable x Patience Index	0.629 (1.374)
Variable x Risk Index	-1.231 (1.523)
Variable x Altruism (0-10)	-0.033 (0.702)
Variable x Extraversion (1-5)	2.684 (1.438)*
Variable x Agreeableness (1-5)	2.366 (2.458)
Variable x Conscientiousness (1-5)	-1.884 (3.897)
Variable x Neuroticism (1-5)	0.232 (1.804)
Variable x Openness (1-5)	-0.847 (1.562)
Variable x Enjoy Competition (1-5)	0.757 (1.265)
Variable x Positive Parenting (1-5)	0.228 (1.225)
Variable x Grit (1-5)	0.432 (2.805)
Variable x Belief on rel. performance (0 - 1)	-4.534 (4.878)
Tournament x Productivity (resid.)	0.424 (0.256)*
Tournament x Female (=1)	0.792 (3.441)
Tournament x Rel. age grade mean	-1.016 (1.361)
Tournament x Grade (9-13)	1.327 (1.628)
Tournament x IQ (Raven 0-10)	0.050 (0.977)
Tournament x Grade Math	-1.758 (1.107)
Tournament x Grade German	2.152 (1.306)*
Tournament x Number of siblings	1.262 (1.176)
Tournament x Composite SES Index	-0.229 (1.094)
Tournament x Patience Index	0.120 (1.860)
Tournament x Risk Index	-2.945 (1.597)*
Tournament x Altruism (0-10)	-0.723 (0.593)
Tournament x Extraversion (1-5)	3.086 (1.555)**
Tournament x Agreeableness (1-5)	1.472 (2.333)
Tournament x Conscientiousness (1-5)	-6.196 (3.542)*
Tournament x Neuroticism (1-5)	-0.499 (1.974)
Tournament x Openness (1-5)	1.192 (1.708)
Tournament x Enjoy Competition (1-5)	0.960 (1.298)
Tournament x Positive Parenting (1-5)	-0.432 (1.427)
Tournament x Grit (1-5)	2.804 (2.746)
Tournament x Belief on rel. performance (0 - 1)	-4.189 (4.607)
Constant	70.368 (19.929)***
Num.Obs.	1933
R2 Adj.	0.396

*Note:*

The table shows OLS regressions for interaction effects between payments and characteristics on performance (Part 2). Standard errors (in parentheses) clustered on the session level.

Table A2: Productivity: Interaction effects between payments and characteristics on performance (Part 2)

	Using Variable as base
	(1)
Fixed	5.084 (24.515)
Tournament	-5.943 (20.716)
Productivity (resid.)	2.064 (0.209)***
Female (=1)	3.485 (1.409)**
Age rel. to grade mean	-1.778 (0.801)**
Grade (9-13)	2.794 (0.865)***
IQ (Raven 0-10)	0.689 (0.414)*
Grade Math	2.169 (0.534)***
Grade German	-0.584 (0.951)
Number of siblings	-0.250 (0.546)
Composite SES Index	-0.580 (0.489)
Patience Index	0.373 (0.725)
Risk Index	-0.121 (0.643)
Altruism (0-10)	0.485 (0.290)*
Extraversion (1-5)	-0.184 (0.989)
Agreeableness (1-5)	0.942 (1.245)
Conscientiousness (1-5)	2.251 (2.204)
Neuroticism (1-5)	-0.024 (0.826)
Openness (1-5)	-1.047 (1.086)
Enjoy Competition (1-5)	1.438 (0.663)**
Positive Parenting (1-5)	-0.923 (0.544)*
Grit (1-5)	-2.090 (1.412)
Belief on rel. performance (0 - 1)	21.759 (2.870)***
Fixed x Productivity (resid.)	-0.353 (0.286)
Fixed x Female (=1)	-1.871 (3.296)
Fixed x Rel. age grade mean	1.202 (1.071)
Fixed x Grade (9-13)	-0.478 (1.632)
Fixed x IQ (Raven 0-10)	0.219 (0.755)
Fixed x Grade Math	1.063 (0.950)
Fixed x Grade German	-0.599 (1.265)
Fixed x Number of siblings	-0.246 (1.078)
Fixed x Composite SES Index	0.783 (0.906)
Fixed x Patience Index	-0.629 (1.374)
Fixed x Risk Index	1.231 (1.523)
Fixed x Altruism (0-10)	0.033 (0.702)
Fixed x Extraversion (1-5)	-2.684 (1.438)*
Fixed x Agreeableness (1-5)	-2.366 (2.458)
Fixed x Conscientiousness (1-5)	1.884 (3.897)
Fixed x Neuroticism (1-5)	-0.232 (1.804)
Fixed x Openness (1-5)	0.847 (1.562)
Fixed x Enjoy Competition (1-5)	-0.757 (1.265)
Fixed x Positive Parenting (1-5)	-0.228 (1.225)
Fixed x Grit (1-5)	-0.432 (2.805)
Fixed x Belief on rel. performance (0 - 1)	4.534 (4.878)
Tournament x Productivity (resid.)	0.071 (0.265)
Tournament x Female (=1)	-1.079 (2.666)
Tournament x Rel. age grade mean	0.186 (1.143)
Tournament x Grade (9-13)	0.848 (0.941)
Tournament x IQ (Raven 0-10)	0.269 (0.614)
Tournament x Grade Math	-0.695 (1.003)
Tournament x Grade German	1.553 (1.379)
Tournament x Number of siblings	1.016 (0.832)
Tournament x Composite SES Index	0.554 (0.991)
Tournament x Patience Index	-0.508 (1.216)
Tournament x Risk Index	-1.713 (1.300)
Tournament x Altruism (0-10)	-0.689 (0.421)
Tournament x Extraversion (1-5)	0.403 (1.454)
Tournament x Agreeableness (1-5)	-0.894 (1.907)
Tournament x Conscientiousness (1-5)	-4.312 (3.361)
Tournament x Neuroticism (1-5)	-0.730 (1.402)
Tournament x Openness (1-5)	2.039 (1.538)
Tournament x Enjoy Competition (1-5)	0.203 (1.009)
Tournament x Positive Parenting (1-5)	-0.660 (0.930)
Tournament x Grit (1-5)	2.373 (2.291)
Tournament x Belief on rel. performance (0 - 1)	0.344 (3.499)
Constant	65.283 (14.699)***
Num.Obs.	1933
R2 Adj.	0.396

*Note:*

The table shows OLS regressions for interaction effects between payments and characteristics on performance (Part 2). Standard errors (in parentheses) clustered on the session level.

Table A3: Productivity: Interaction effects between payments and characteristics on performance (Part 2)

	Using Fixed as base
	(1)
Variable	-19.322 (31.922)
Tournament	-15.226 (26.909)
Productivity (resid.)	1.253 (0.571)**
Female (=1)	-0.180 (5.200)
Age rel. to grade mean	0.973 (1.343)
Grade (9-13)	1.251 (1.951)
IQ (Raven 0-10)	1.064 (1.430)
Grade Math	1.871 (1.371)
Grade German	-3.301 (1.488)**
Number of siblings	-1.211 (1.084)
Composite SES Index	0.512 (1.381)
Patience Index	-2.860 (2.175)
Risk Index	-0.533 (2.701)
Altruism (0-10)	0.977 (1.033)
Extraversion (1-5)	-4.048 (2.189)*
Agreeableness (1-5)	-2.951 (3.832)
Conscientiousness (1-5)	6.322 (5.324)
Neuroticism (1-5)	-0.945 (3.534)
Openness (1-5)	0.848 (2.674)
Enjoy Competition (1-5)	-0.708 (1.326)
Positive Parenting (1-5)	-2.635 (2.190)
Grit (1-5)	-2.178 (5.008)
Belief on rel. performance (0 - 1)	26.595 (7.938)***
Variable x Productivity (resid.)	1.249 (0.525)**
Variable x Female (=1)	4.354 (5.478)
Variable x Rel. age grade mean	-1.340 (1.428)
Variable x Grade (9-13)	0.543 (2.196)
Variable x IQ (Raven 0-10)	-0.281 (1.341)
Variable x Grade Math	0.086 (1.556)
Variable x Grade German	2.795 (1.760)
Variable x Number of siblings	1.223 (1.035)
Variable x Composite SES Index	-1.423 (1.516)
Variable x Patience Index	2.872 (2.279)
Variable x Risk Index	0.737 (2.810)
Variable x Altruism (0-10)	-0.677 (1.212)
Variable x Extraversion (1-5)	3.979 (2.141)*
Variable x Agreeableness (1-5)	5.639 (4.392)
Variable x Conscientiousness (1-5)	-3.377 (6.126)
Variable x Neuroticism (1-5)	1.460 (3.643)
Variable x Openness (1-5)	-2.671 (3.003)
Variable x Enjoy Competition (1-5)	2.190 (1.158)*
Variable x Positive Parenting (1-5)	1.517 (2.046)
Variable x Grit (1-5)	-1.474 (5.547)
Variable x Belief on rel. performance (0 - 1)	-6.643 (8.581)
Tournament x Productivity (resid.)	1.041 (0.461)**
Tournament x Female (=1)	3.314 (5.989)
Tournament x Rel. age grade mean	-1.517 (1.496)
Tournament x Grade (9-13)	1.148 (2.212)
Tournament x IQ (Raven 0-10)	0.228 (2.024)
Tournament x Grade Math	-0.904 (1.658)
Tournament x Grade German	5.004 (2.047)**
Tournament x Number of siblings	1.891 (1.701)
Tournament x Composite SES Index	1.070 (1.573)
Tournament x Patience Index	3.457 (2.995)
Tournament x Risk Index	-1.838 (2.702)
Tournament x Altruism (0-10)	-1.530 (0.904)*
Tournament x Extraversion (1-5)	5.769 (2.410)**
Tournament x Agreeableness (1-5)	5.183 (4.269)
Tournament x Conscientiousness (1-5)	-12.914 (6.216)**
Tournament x Neuroticism (1-5)	-0.326 (3.460)
Tournament x Openness (1-5)	-1.026 (3.162)
Tournament x Enjoy Competition (1-5)	0.413 (2.072)
Tournament x Positive Parenting (1-5)	2.325 (2.243)
Tournament x Grit (1-5)	4.908 (5.819)
Tournament x Belief on rel. performance (0 - 1)	-7.613 (8.124)
Constant	94.784 (24.579)***
Num.Obs.	950
R2 Adj.	0.488

*Note:*

The table shows OLS regressions for interaction effects between payments and characteristics on performance (Part 2). Standard errors (in parentheses) clustered on the session level.

Table A4: Productivity: Interaction effects between payments and characteristics on performance (Part 2)

	Using Variable as base
	(1)
Fixed	19.322 (31.922)
Tournament	4.096 (24.474)
Productivity (resid.)	2.501 (0.227)***
Female (=1)	4.173 (1.569)***
Age rel. to grade mean	-0.366 (0.988)
Grade (9-13)	1.794 (0.843)**
IQ (Raven 0-10)	0.783 (0.496)
Grade Math	1.958 (0.752)***
Grade German	-0.506 (0.957)
Number of siblings	0.012 (0.715)
Composite SES Index	-0.911 (0.490)*
Patience Index	0.011 (0.762)
Risk Index	0.204 (0.786)
Altruism (0-10)	0.300 (0.372)
Extraversion (1-5)	-0.068 (1.401)
Agreeableness (1-5)	2.687 (1.540)*
Conscientiousness (1-5)	2.944 (2.290)
Neuroticism (1-5)	0.516 (1.043)
Openness (1-5)	-1.823 (1.297)
Enjoy Competition (1-5)	1.482 (0.681)**
Positive Parenting (1-5)	-1.118 (0.687)
Grit (1-5)	-3.653 (1.495)**
Belief on rel. performance (0 - 1)	19.951 (3.910)***
Fixed x Productivity (resid.)	-1.249 (0.525)**
Fixed x Female (=1)	-4.354 (5.478)
Fixed x Rel. age grade mean	1.340 (1.428)
Fixed x Grade (9-13)	-0.543 (2.196)
Fixed x IQ (Raven 0-10)	0.281 (1.341)
Fixed x Grade Math	-0.086 (1.556)
Fixed x Grade German	-2.795 (1.760)
Fixed x Number of siblings	-1.223 (1.035)
Fixed x Composite SES Index	1.423 (1.516)
Fixed x Patience Index	-2.872 (2.279)
Fixed x Risk Index	-0.737 (2.810)
Fixed x Altruism (0-10)	0.677 (1.212)
Fixed x Extraversion (1-5)	-3.979 (2.141)*
Fixed x Agreeableness (1-5)	-5.639 (4.392)
Fixed x Conscientiousness (1-5)	3.377 (6.126)
Fixed x Neuroticism (1-5)	-1.460 (3.643)
Fixed x Openness (1-5)	2.671 (3.003)
Fixed x Enjoy Competition (1-5)	-2.190 (1.158)*
Fixed x Positive Parenting (1-5)	-1.517 (2.046)
Fixed x Grit (1-5)	1.474 (5.547)
Fixed x Belief on rel. performance (0 - 1)	6.643 (8.581)
Tournament x Productivity (resid.)	-0.207 (0.393)
Tournament x Female (=1)	-1.040 (4.295)
Tournament x Rel. age grade mean	-0.177 (1.321)
Tournament x Grade (9-13)	0.605 (1.443)
Tournament x IQ (Raven 0-10)	0.510 (1.017)
Tournament x Grade Math	-0.991 (1.417)
Tournament x Grade German	2.209 (1.700)
Tournament x Number of siblings	0.668 (1.420)
Tournament x Composite SES Index	2.493 (1.174)**
Tournament x Patience Index	0.585 (1.633)
Tournament x Risk Index	-2.575 (1.515)*
Tournament x Altruism (0-10)	-0.853 (0.610)
Tournament x Extraversion (1-5)	1.789 (1.813)
Tournament x Agreeableness (1-5)	-0.456 (2.184)
Tournament x Conscientiousness (1-5)	-9.537 (4.115)**
Tournament x Neuroticism (1-5)	-1.786 (2.194)
Tournament x Openness (1-5)	1.645 (1.812)
Tournament x Enjoy Competition (1-5)	-1.778 (1.654)
Tournament x Positive Parenting (1-5)	0.808 (1.255)
Tournament x Grit (1-5)	6.382 (2.771)**
Tournament x Belief on rel. performance (0 - 1)	-0.970 (4.554)
Constant	75.462 (15.122)***
Num.Obs.	950
R2 Adj.	0.488

*Note:*

The table shows OLS regressions for interaction effects between payments and characteristics on performance (Part 2). Standard errors (in parentheses) clustered on the session level.

Table A5: Productivity: Interaction effects between payments and characteristics on performance (Part 2)

	Logit						Ordered Logit
	Payments						F - V - T
	F vs. V or T	V vs. F or T	V vs. F	T vs. V or F	T vs. V	T vs. F	
(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	
<b>Skills</b>							
Grade German	-0.226*** (0.086)	0.099 (0.101)	0.222** (0.102)	0.099 (0.105)	0.038 (0.128)	0.191* (0.098)	0.171** (0.068)
Grade Math	-0.107 (0.073)	-0.099* (0.060)	0.041 (0.078)	0.258*** (0.093)	0.231** (0.097)	0.315*** (0.109)	0.161** (0.065)
IQ (Raven 0-10)	-0.027 (0.065)	-0.021 (0.049)	-0.004 (0.066)	0.053 (0.058)	0.074 (0.063)	0.056 (0.086)	0.043 (0.049)
Productivity (resid.)	-0.101*** (0.016)	0.014 (0.011)	0.091*** (0.015)	0.097*** (0.025)	0.066*** (0.025)	0.147*** (0.038)	0.101*** (0.015)
<b>Demographics</b>							
Age rel. to grade mean	-0.016 (0.146)	0.057 (0.103)	0.034 (0.147)	-0.097 (0.114)	-0.116 (0.129)	-0.058 (0.175)	-0.031 (0.102)
Female (=1)	-0.400** (0.181)	0.687*** (0.151)	0.571*** (0.168)	-0.585** (0.258)	-0.747*** (0.258)	-0.137 (0.322)	-0.094 (0.202)
Grade (9-13)	0.120 (0.145)	-0.198* (0.103)	-0.165 (0.144)	0.123 (0.095)	0.183* (0.105)	-0.043 (0.154)	0.021 (0.096)
Number of siblings	0.033 (0.069)	-0.029 (0.047)	-0.032 (0.067)	0.028 (0.065)	0.038 (0.063)	-0.011 (0.114)	-0.009 (0.051)
Positive Parenting (0-5)	0.162 (0.107)	-0.044 (0.074)	-0.146 (0.106)	-0.113 (0.100)	-0.061 (0.100)	-0.228 (0.147)	-0.130 (0.083)
Composite SES Index	-0.024 (0.096)	0.021 (0.068)	0.016 (0.095)	-0.026 (0.074)	-0.014 (0.085)	-0.072 (0.117)	0.011 (0.072)
<b>Personality Traits, Economic Preferences, and Beliefs</b>							
Altruism (0-10)	-0.033 (0.037)	0.047 (0.037)	0.051 (0.039)	-0.013 (0.049)	-0.032 (0.052)	0.039 (0.059)	0.005 (0.037)
Extraversion (0-5)	0.264** (0.119)	-0.170* (0.101)	-0.284** (0.126)	-0.007 (0.102)	0.040 (0.109)	-0.153 (0.158)	-0.154* (0.084)
Agreeableness (0-5)	0.091 (0.153)	0.094 (0.141)	-0.017 (0.161)	-0.222 (0.174)	-0.209 (0.178)	-0.327 (0.212)	-0.134 (0.130)
Conscientiousness (0-5)	-0.188 (0.252)	0.261 (0.235)	0.210 (0.265)	-0.207 (0.244)	-0.314 (0.263)	0.301 (0.312)	-0.043 (0.185)
Neuroticism (0-5)	0.268** (0.126)	-0.168 (0.103)	-0.289** (0.130)	-0.027 (0.155)	0.044 (0.147)	-0.160 (0.203)	-0.163 (0.118)
Openness (0-5)	0.139 (0.124)	-0.085 (0.090)	-0.114 (0.117)	0.024 (0.146)	0.124 (0.153)	-0.175 (0.198)	-0.055 (0.114)
Enjoy Competition (1-5)	-0.408*** (0.071)	-0.108 (0.087)	0.253*** (0.082)	0.601*** (0.094)	0.516*** (0.104)	0.801*** (0.117)	0.480*** (0.062)
Grit (1-5)	0.175 (0.217)	-0.125 (0.178)	-0.209 (0.233)	0.033 (0.212)	0.151 (0.225)	-0.348 (0.251)	-0.035 (0.186)
Belief on rel. performance (0-1)	-0.726** (0.287)	-0.437** (0.218)	0.376 (0.302)	1.285*** (0.270)	1.133*** (0.273)	1.366*** (0.383)	1.022*** (0.230)
Patience Index	-0.181** (0.074)	0.046 (0.086)	0.151* (0.079)	0.110 (0.136)	0.026 (0.143)	0.191 (0.142)	0.140 (0.090)
Risk Taking Index	-0.244** (0.108)	-0.055 (0.098)	0.154 (0.106)	0.325** (0.132)	0.308** (0.144)	0.420** (0.182)	0.301*** (0.098)
Constant	-1.964 (2.094)	2.872* (1.675)	2.679 (2.234)	-4.876*** (1.081)	-5.257*** (1.450)	-1.777 (1.860)	
Num.Obs.	950	950	693	950	715	492	950
RMSE	0.41	0.48	0.45	0.39	0.43	0.41	1.84

\* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

Note: Models (I) - (VII) are logit regressions. Coefficients represent odds ratios (Coefficients above (below) 1 represent a positive (negative) association between the predictor and the outcome). Model (I) compares choosing the Fixed (F) over the Variable (V) or Tournament (T) payment. Model (II) compares choosing the variable over the fixed or tournament payment. Model (III) compares choosing the Tournament over the Variable or Fixed payment. Model (IV) compares choosing the Fixed over the Variable payment. Model (V) compares choosing the Variable over the Tournament payment. Model (VI) compares choosing the Fixed over the Tournament payment. Model (VII) is a multinomial logit regression (f - v - t); Clustered standard errors on the session level.

Table A6: Choice regressions



	Logit						Ordered Logit
	Payments						F - V - T
	F vs. V or T	V vs. F or T	V vs. F	T vs. V or F	T vs. V	T vs. F	
(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	
<b>Skills</b>							
Grade German	-0.237*** (0.088)	0.105 (0.101)	0.222** (0.101)	0.099 (0.105)	0.034 (0.129)	0.199** (0.099)	0.175*** (0.067)
Grade Math	-0.107 (0.073)	-0.099 (0.063)	0.045 (0.084)	0.258*** (0.093)	0.231** (0.098)	0.315*** (0.108)	0.162** (0.063)
IQ (Raven 0-10)	-0.034 (0.064)	-0.019 (0.049)	0.001 (0.065)	0.053 (0.058)	0.073 (0.063)	0.060 (0.087)	0.045 (0.049)
Productivity (resid.)	-0.100*** (0.017)	0.013 (0.011)	0.092*** (0.016)	0.097*** (0.025)	0.066*** (0.025)	0.146*** (0.038)	0.101*** (0.015)
<b>Demographics</b>							
Age rel. to grade mean	0.012 (0.158)	0.034 (0.100)	-0.012 (0.157)	-0.097 (0.114)	-0.113 (0.124)	-0.077 (0.188)	-0.048 (0.115)
Female (=1)	-0.376** (0.188)	0.676*** (0.154)	0.565*** (0.172)	-0.585** (0.259)	-0.740*** (0.260)	-0.151 (0.328)	-0.103 (0.203)
Grade (9-13)	0.094 (0.141)	-0.174* (0.096)	-0.125 (0.138)	0.123 (0.098)	0.174 (0.108)	-0.042 (0.150)	0.045 (0.101)
Number of siblings	0.042 (0.069)	-0.037 (0.048)	-0.042 (0.066)	0.028 (0.067)	0.040 (0.064)	-0.020 (0.114)	-0.015 (0.052)
Positive Parenting (0-5)	0.168 (0.107)	-0.049 (0.074)	-0.158 (0.107)	-0.113 (0.100)	-0.061 (0.100)	-0.231 (0.147)	-0.133 (0.082)
Composite SES Index	-0.020 (0.094)	0.018 (0.065)	0.014 (0.091)	-0.026 (0.074)	-0.010 (0.086)	-0.069 (0.114)	0.007 (0.071)
<b>Personality Traits, Economic Preferences, and Beliefs</b>							
Altruism (0-10)	-0.038 (0.037)	0.050 (0.037)	0.056 (0.039)	-0.013 (0.049)	-0.032 (0.052)	0.043 (0.061)	0.006 (0.038)
Extraversion (0-5)	0.276** (0.122)	-0.175* (0.101)	-0.294** (0.127)	-0.007 (0.102)	0.043 (0.110)	-0.159 (0.159)	-0.158* (0.084)
Agreeableness (0-5)	0.097 (0.155)	0.090 (0.138)	-0.013 (0.163)	-0.222 (0.173)	-0.210 (0.177)	-0.326 (0.214)	-0.136 (0.131)
Conscientiousness (0-5)	-0.185 (0.263)	0.253 (0.239)	0.198 (0.275)	-0.207 (0.244)	-0.309 (0.263)	0.306 (0.305)	-0.052 (0.189)
Neuroticism (0-5)	0.271** (0.125)	-0.170* (0.102)	-0.298** (0.128)	-0.027 (0.155)	0.045 (0.148)	-0.154 (0.200)	-0.166 (0.119)
Openness (0-5)	0.123 (0.123)	-0.068 (0.091)	-0.105 (0.118)	0.024 (0.142)	0.116 (0.148)	-0.167 (0.194)	-0.042 (0.110)
Enjoy Competition (1-5)	-0.412*** (0.073)	-0.108 (0.086)	0.265*** (0.081)	0.601*** (0.094)	0.515*** (0.103)	0.800*** (0.118)	0.484*** (0.063)
Grit (1-5)	0.168 (0.220)	-0.127 (0.181)	-0.208 (0.240)	0.032 (0.211)	0.152 (0.226)	-0.355 (0.250)	-0.036 (0.187)
Belief on rel. performance (0-1)	-0.755*** (0.272)	-0.414** (0.208)	0.414 (0.276)	1.285*** (0.268)	1.123*** (0.268)	1.377*** (0.378)	1.045*** (0.225)
Patience Index	-0.184** (0.075)	0.047 (0.085)	0.159** (0.080)	0.110 (0.136)	0.026 (0.143)	0.184 (0.143)	0.142 (0.091)
Risk Taking Index	-0.258** (0.108)	-0.049 (0.098)	0.171 (0.105)	0.326** (0.132)	0.306** (0.144)	0.425** (0.181)	0.307*** (0.098)
<b>Time Difference between Sessions (in days)</b>							
Time Difference	0.026 (0.018)	-0.020 (0.016)	-0.030 (0.020)	0.000 (0.012)	0.006 (0.014)	-0.013 (0.020)	-0.016 (0.012)
Constant	-1.906 (1.975)	2.815* (1.538)	2.593 (2.112)	-4.877*** (1.073)	-5.203*** (1.425)	-1.699 (1.778)	
Num.Obs.	950	950	693	950	715	492	950
RMSE	0.40	0.48	0.45	0.39	0.43	0.41	1.85

\* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

Note: Models (I) - (VII) are logit regressions. Coefficients represent odds ratios (Coefficients above (below) 1 represent a positive (negative) association between the predictor and the outcome). Model (I) compares choosing the Fixed (F) over the Variable (V) or Tournament (T) payment. Model (II) compares choosing the variable over the fixed or tournament payment. Model (III) compares choosing the Tournament over the Variable or Fixed payment. Model (IV) compares choosing the Fixed over the Variable payment. Model (V) compares choosing the Variable over the Tournament payment. Model (VI) compares choosing the Fixed over the Tournament payment. Model (VII) is a multinomial logit regression (f - v - t); Clustered standard errors on the session level.

Table A7: Choice regressions

	Belief on rel. performance (0-1)	Overconfidence (-1 - 1)
	(I)	(II)
<b>Skills</b>		
Grade German	0.000 (0.007)	-0.015*** (0.005)
Grade Math	0.009* (0.005)	-0.006 (0.004)
IQ (Raven 0-10)	-0.001 (0.003)	-0.003 (0.002)
Productivity (resid.)	0.000 (0.001)	-0.042*** (0.002)
<b>Demographics</b>		
Age rel. to grade mean	0.007 (0.008)	0.028*** (0.006)
Female (=1)	-0.027** (0.013)	-0.037*** (0.010)
Grade (9-13)	0.009 (0.006)	-0.004 (0.007)
Number of siblings	0.001 (0.005)	0.004 (0.004)
Positive Parenting (1-5)	0.011* (0.006)	0.039*** (0.006)
Composite SES Index	0.015*** (0.006)	0.012** (0.005)
<b>Personality Traits, Economic Preferences, and Beliefs</b>		
Altruism (0-10)	0.001 (0.003)	0.000 (0.002)
Extraversion (1-5)	-0.004 (0.007)	0.018*** (0.006)
Agreeableness (1-5)	0.001 (0.011)	0.020** (0.008)
Conscientiousness (1-5)	-0.003 (0.018)	-0.013 (0.015)
Neuroticism (1-5)	-0.014 (0.010)	0.002 (0.008)
Openness (1-5)	0.000 (0.011)	-0.011 (0.007)
Enjoy Competition (1-5)	0.000 (0.006)	-0.018*** (0.005)
Grit (1-5)	0.002 (0.016)	-0.005 (0.010)
Patience Index	-0.002 (0.008)	-0.016*** (0.006)
Risk Taking Index	0.020*** (0.007)	0.007 (0.006)
Constant	0.433*** (0.104)	0.111 (0.106)
<b>Num.Obs.</b>	1933	1933
<b>R2 Adj.</b>	0.008	0.635

\* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

*Note:*

Overconfidence equals the spread between the belief and actual relative performance (both in terms of rank percentiles. Clustered standard errors on the session level

Table A8: Determinants of beliefs on rel. self-assessment and overconfidence

	Exogenous			Payments			Endogenous		
	Fixed	Variable	Tournament	Fixed	Tournament	Variable	Fixed	Variable	Tournament
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)
<b>Skills</b>									
Grade German (bottom quartile)	3.886 (2.295)*	0.479 (2.482)	-3.571 (2.511)	7.051 (3.194)**	1.641 (1.835)	1.191 (2.082)			
Grade Math (bottom quartile)	-16.174 (2.401)***	-6.912 (3.066)**	0.642 (4.055)	-3.456 (3.945)	-6.843 (2.344)***	-8.592 (5.553)			
IQ (Raven 0-10) (bottom quartile)	-0.860 (3.158)	-1.815 (2.347)	2.180 (2.348)	-4.420 (5.796)	-1.329 (2.538)	-0.744 (2.778)			
Productivity (resid.)	1.681 (0.239)***	1.623 (0.311)***	1.839 (0.344)***	1.162 (0.580)**	2.471 (0.216)***	2.445 (0.330)***			
<b>Demographics</b>									
Rel. age grade mean (bottom quartile)	6.622 (2.161)***	8.662 (2.700)***	0.035 (2.948)	5.628 (3.921)	3.338 (1.946)*	-1.330 (2.446)			
Female (=1)	3.968 (2.283)*	2.377 (2.186)	-0.584 (2.054)	0.203 (4.150)	2.846 (1.525)*	2.373 (2.989)			
Grade (9-13)	4.848 (0.944)***	5.338 (1.580)***	4.460 (1.097)***	2.305 (1.662)	2.372 (0.907)***	2.433 (1.422)*			
Number of siblings (bottom quartile)	-5.825 (3.471)*	1.380 (3.862)	-3.513 (2.473)	-2.331 (3.035)	0.710 (1.695)	-5.609 (2.629)**			
Positive Parenting (1-5) (bottom quartile)	-1.517 (2.566)	2.029 (2.665)	5.306 (2.101)**	3.495 (4.120)	0.455 (1.484)	2.933 (1.943)			
SES Index (bottom quartile)	-1.411 (2.339)	0.457 (2.008)	3.248 (2.564)	-6.580 (4.192)	1.661 (1.387)	0.037 (2.164)			
<b>Personality Traits, Economic Preferences, and Beliefs</b>									
Altruism (0-10) (bottom quartile)	1.080 (2.763)	-2.269 (2.341)	1.668 (2.325)	-0.095 (4.252)	-0.834 (1.638)	0.728 (1.963)			
Extraversion (1-5) (bottom quartile)	2.671 (2.589)	0.540 (2.607)	4.342 (2.489)*	7.842 (3.397)**	1.850 (2.259)	-4.949 (2.276)**			
Agreeableness (1-5) (bottom quartile)	1.472 (3.401)	-0.455 (2.265)	0.858 (2.825)	1.761 (3.663)	-3.588 (1.859)*	-2.616 (2.049)			
Conscientiousness (1-5) (bottom quartile)	-0.234 (1.982)	2.908 (2.842)	-4.967 (2.656)*	-4.695 (3.645)	-1.916 (1.900)	4.799 (2.323)**			
Neuroticism (1-5) (bottom quartile)	-0.419 (2.933)	1.303 (1.917)	0.932 (2.518)	2.440 (4.370)	-0.849 (2.084)	2.354 (2.506)			
Openness (1-5) (bottom quartile)	-6.253 (2.103)***	0.129 (3.154)	-0.934 (2.349)	-7.105 (4.034)*	0.051 (1.613)	0.058 (1.626)			
Enjoy Competition (1-5) (bottom quartile)	-1.132 (2.985)	-3.889 (3.068)	-2.387 (3.173)	-0.022 (3.229)	0.444 (1.693)	0.959 (3.017)			
Grit (1-5) (bottom quartile)	1.093 (2.981)	0.377 (2.518)	-0.829 (2.276)	-2.542 (4.754)	3.474 (1.850)*	-3.567 (1.897)*			
Belief on rel. performance (0-1) (bottom quartile)	-8.240 (2.759)***	-10.398 (2.740)***	-4.849 (2.649)*	-10.487 (4.211)**	-7.643 (1.980)***	-8.537 (1.989)***			
Patience Index (bottom quartile)	-1.123 (2.023)	0.070 (1.976)	-1.933 (2.478)	4.852 (3.655)	0.709 (1.712)	-0.295 (2.176)			
Risk Taking Index (bottom quartile)	-0.851 (2.166)	2.344 (2.732)	0.052 (2.258)	3.908 (3.406)	-1.448 (1.480)	5.457 (2.509)**			
Constant	67.055 (12.033)***	64.163 (17.264)***	73.710 (12.757)***	80.283 (21.495)***	97.800 (9.283)***	102.209 (16.188)***			
Num.Obs.	331	327	325	235	458	257			
R2 Adj.	0.284	0.292	0.269	0.109	0.435	0.441			

Note: Table shows OLS regressions for bottom quartile interaction effects on performance in part 2 split by treatment and payment. Standard errors (in parentheses) clustered on the session level.

Table A9: Productivity: Interaction effects for the variable incentive scheme by treatment - Bottom quartile

	Exogenous			Payment			Endogenous		
	Fixed	Variable	Tournament	Fixed	Variable	Tournament	Fixed	Variable	Tournament
	(I)	(II)	(III)	(IV)	(V)	(VI)	(IV)	(V)	(VI)
<b>Skills</b>									
Grade German (top quartile)	2.149 (3.396)	4.666 (4.817)	7.447 (3.831)*	-8.658 (8.856)	0.598 (2.845)	6.527 (5.983)	-8.658 (8.856)	0.598 (2.845)	6.527 (5.983)
Grade Math (top quartile)	5.945 (2.364)**	4.210 (3.340)	-0.421 (3.363)	3.840 (6.474)	7.299 (2.834)**	-0.274 (2.771)	3.840 (6.474)	7.299 (2.834)**	-0.274 (2.771)
IQ (Raven 0-10) (top quartile)	8.065 (2.068)***	2.328 (2.476)	1.439 (2.450)	1.653 (4.237)	2.055 (2.038)	4.798 (2.340)**	1.653 (4.237)	2.055 (2.038)	4.798 (2.340)**
Productivity (resid.)	1.632 (0.240)***	1.597 (0.246)***	1.908 (0.294)***	1.300 (0.610)**	2.425 (0.217)***	2.156 (0.315)***	1.300 (0.610)**	2.425 (0.217)***	2.156 (0.315)***
<b>Demographics</b>									
Rel. age grade mean (top quartile)	-5.965 (1.849)***	-5.793 (2.942)**	-7.389 (2.103)***	6.783 (3.904)*	-1.218 (1.669)	-2.880 (2.617)	6.783 (3.904)*	-1.218 (1.669)	-2.880 (2.617)
Female (=1)	3.956 (2.256)*	2.995 (2.746)	0.714 (2.230)	-0.977 (3.418)	4.709 (1.410)***	1.200 (2.651)	-0.977 (3.418)	4.709 (1.410)***	1.200 (2.651)
Grade (9-13)	3.698 (0.854)***	3.847 (1.632)**	5.180 (1.160)***	1.047 (2.375)	1.951 (0.760)**	2.496 (1.325)*	1.047 (2.375)	1.951 (0.760)**	2.496 (1.325)*
Number of siblings (top quartile)	-0.564 (3.272)	-3.383 (2.071)	0.858 (2.528)	-2.385 (3.096)	-0.619 (2.421)	0.991 (3.679)	-2.385 (3.096)	-0.619 (2.421)	0.991 (3.679)
Positive Parenting (1-5) (top quartile)	0.851 (1.811)	-0.786 (2.143)	-3.310 (2.495)	0.240 (5.447)	-0.718 (1.500)	0.483 (2.322)	0.240 (5.447)	-0.718 (1.500)	0.483 (2.322)
SES Index (top quartile)	-0.161 (1.935)	0.419 (2.545)	2.046 (1.875)	-1.211 (3.627)	-2.952 (1.500)**	2.721 (2.693)	-1.211 (3.627)	-2.952 (1.500)**	2.721 (2.693)
<b>Personality Traits, Economic Preferences, and Beliefs</b>									
Altruism (0-10) (top quartile)	1.040 (2.193)	1.300 (2.224)	-2.342 (1.899)	-3.290 (5.092)	-0.104 (1.844)	-3.338 (3.098)	-3.290 (5.092)	-0.104 (1.844)	-3.338 (3.098)
Extraversion (1-5) (top quartile)	-2.178 (2.954)	0.682 (2.801)	-1.422 (2.876)	0.007 (3.070)	0.211 (2.846)	-2.505 (2.338)	0.007 (3.070)	0.211 (2.846)	-2.505 (2.338)
Agreeableness (1-5) (top quartile)	1.321 (2.448)	-4.905 (2.744)*	-1.532 (2.241)	-3.309 (5.801)	2.164 (2.315)	4.915 (2.228)**	-3.309 (5.801)	2.164 (2.315)	4.915 (2.228)**
Conscientiousness (1-5) (top quartile)	-2.148 (2.535)	-0.257 (2.728)	1.563 (2.881)	6.904 (4.030)*	-0.762 (1.839)	1.952 (3.189)	6.904 (4.030)*	-0.762 (1.839)	1.952 (3.189)
Neuroticism (1-5) (top quartile)	-1.703 (2.281)	-1.044 (2.446)	-0.485 (2.229)	2.523 (3.756)	-0.307 (1.595)	1.467 (2.560)	2.523 (3.756)	-0.307 (1.595)	1.467 (2.560)
Openness (1-5) (top quartile)	-3.782 (3.389)	-2.850 (2.134)	1.373 (2.342)	-2.792 (3.901)	-3.262 (1.943)*	1.693 (1.865)	-2.792 (3.901)	-3.262 (1.943)*	1.693 (1.865)
Enjoy Competition (1-5) (top quartile)	-1.601 (3.557)	1.656 (2.299)	7.172 (2.365)***	-3.183 (4.654)	1.965 (1.504)	0.601 (2.313)	-3.183 (4.654)	1.965 (1.504)	0.601 (2.313)
Grit (1-5) (top quartile)	-1.574 (2.534)	1.372 (2.389)	-1.362 (1.814)	-3.043 (6.028)	0.131 (1.864)	1.134 (3.694)	-3.043 (6.028)	0.131 (1.864)	1.134 (3.694)
Belief on rel. performance (0 - 1) (top quartile)	13.453 (2.738)***	11.274 (1.782)***	11.955 (1.837)***	14.066 (4.232)***	8.441 (1.591)***	10.927 (1.892)***	14.066 (4.232)***	8.441 (1.591)***	10.927 (1.892)***
Patience Index (top quartile)	2.555 (2.138)	0.169 (2.244)	-3.809 (2.040)*	-0.080 (3.993)	0.710 (1.522)	-2.717 (2.073)	-0.080 (3.993)	0.710 (1.522)	-2.717 (2.073)
Risk Taking Index (top quartile)	1.370 (2.113)	-2.459 (2.271)	-2.766 (2.895)	1.484 (3.837)	1.515 (1.798)	-1.652 (2.038)	1.484 (3.837)	1.515 (1.798)	-1.652 (2.038)
Constant	74.065 (10.216)***	79.422 (19.482)***	63.679 (12.583)***	91.875 (27.251)***	97.300 (7.999)***	94.727 (15.407)***	91.875 (27.251)***	97.300 (7.999)***	94.727 (15.407)***
Num. Obs.	331	327	325	235	458	257	235	458	257
R2 Adj.	0.308	0.304	0.345	0.068	0.445	0.475	0.068	0.445	0.475

Note: Table shows OLS regressions for top quartile interaction effects on performance in part 2 split by treatment and payment. Standard errors (in parentheses) clustered on the session level.

Table A10: Productivity: Interaction effects for the variable incentive scheme by treatment - top quartile

	Exogenous	Endogenous
	(I)	(II)
<b>Variable</b>	4.646 (1.209)***	16.461 (1.846)***
<b>Tournament</b>	2.598 (1.401)*	19.214 (2.174)***
<b>Skills</b>		
Grade German	-0.096 (0.682)	-0.981 (0.577)*
Grade Math	2.365 (0.502)***	1.860 (0.500)***
IQ (Raven 0-10)	0.603 (0.392)	1.088 (0.445)**
Productivity (resid.)	1.710 (0.224)***	2.097 (0.293)***
<b>Demographics</b>		
Age rel. to grade mean	-2.945 (0.506)***	-0.339 (0.834)
Female (=1)	2.597 (1.063)**	2.658 (1.583)*
Grade (9-13)	3.833 (0.737)***	1.667 (0.770)**
Number of siblings	0.082 (0.502)	0.022 (0.666)
Positive Parenting (1-5)	-0.926 (0.639)	-1.103 (0.770)
Composite SES Index	-0.456 (0.540)	-0.006 (0.577)
<b>Personality Traits, Economic Preferences, and Beliefs</b>		
Altruism (0-10)	0.226 (0.294)	0.165 (0.301)
Extraversion (1-5)	-0.925 (0.711)	-0.349 (1.127)
Agreeableness (1-5)	-0.673 (1.406)	1.134 (0.845)
Conscientiousness (1-5)	1.550 (1.892)	1.641 (1.766)
Neuroticism (1-5)	-0.449 (0.751)	-0.069 (1.003)
Openness (1-5)	0.742 (0.925)	-0.815 (0.951)
Enjoy Competition (1-5)	1.215 (0.622)*	0.609 (0.561)
Grit (1-5)	-1.572 (1.100)	-1.465 (1.397)
Belief on rel. performance (perc)	22.720 (3.015)***	21.678 (3.295)***
Patience Index	0.433 (0.745)	-0.761 (0.601)
Risk Taking Index	-0.366 (0.859)	-0.438 (0.743)
<b>Constant</b>	52.569 (12.458)***	67.232 (12.568)***
<b>Num.Obs.</b>	983	950
<b>R2 Adj.</b>	0.356	0.471

Note: Table shows OLS regressions on performance in part 2 split by treatment. Standard errors (in parentheses) clustered on the session level.

Table A11: Productivity by treatment

	Exogenous			Endogenous		
	Payment					
	Fixed	Variable	Tournament	Fixed	Variable	Tournament
(I)	(II)	(III)	(IV)	(V)	(VI)	
<b>Skills</b>						
Grade German	-0.540 (0.933)	-0.381 (1.742)	0.405 (1.216)	-3.323 (1.514)**	-0.500 (0.948)	1.687 (1.202)
Grade Math	4.184 (0.908)***	2.220 (0.856)***	1.399 (0.941)	2.004 (1.409)	1.958 (0.747)***	0.971 (1.100)
IQ (Raven 0-10)	0.143 (0.539)	0.434 (0.639)	0.754 (0.571)	1.099 (1.486)	0.786 (0.489)	1.290 (0.876)
Productivity (resid.)	1.720 (0.221)***	1.571 (0.269)***	1.888 (0.314)***	1.246 (0.580)**	2.502 (0.225)***	2.295 (0.300)***
<b>Demographics</b>						
Age rel. to grade mean	-2.876 (0.784)***	-3.610 (1.274)***	-1.854 (1.145)	0.297 (1.276)	-0.379 (0.972)	-0.542 (1.116)
Female (=1)	4.626 (2.565)*	2.671 (2.492)	1.532 (2.248)	-0.288 (5.247)	4.166 (1.556)***	3.166 (3.456)
Grade (9-13)	4.132 (0.842)***	3.662 (1.558)**	4.651 (0.936)***	1.211 (1.877)	1.830 (0.824)**	2.393 (1.516)
Number of siblings	-0.229 (1.323)	-0.975 (0.825)	0.686 (0.745)	-1.434 (1.139)	0.009 (0.711)	0.688 (1.401)
Positive Parenting (0-5)	0.920 (1.383)	-0.976 (1.117)	-2.202 (1.057)**	-2.924 (2.311)	-1.119 (0.680)	-0.311 (1.117)
Composite SES Index	-0.159 (0.840)	-0.674 (1.044)	-1.059 (1.100)	0.637 (1.373)	-0.919 (0.488)*	1.586 (1.221)
<b>Personality Traits, Economic Preferences, and Beliefs</b>						
Altruism (0-10)	-0.285 (0.419)	0.755 (0.495)	0.064 (0.589)	1.094 (1.039)	0.300 (0.368)	-0.555 (0.416)
Extraversion (0-5)	-1.548 (1.302)	0.056 (1.303)	-0.827 (1.578)	-3.940 (2.150)*	-0.075 (1.396)	1.727 (1.256)
Agreeableness (0-5)	1.398 (2.816)	-1.402 (1.948)	-1.589 (1.930)	-2.725 (3.920)	2.690 (1.528)*	2.237 (1.454)
Conscientiousness (0-5)	0.005 (2.431)	0.734 (4.167)	1.849 (2.735)	6.629 (5.314)	2.929 (2.260)	-6.585 (3.302)**
Neuroticism (0-5)	-0.432 (1.244)	-0.552 (1.322)	-0.964 (1.504)	-0.825 (3.518)	0.510 (1.039)	-1.274 (1.832)
Openness (0-5)	1.389 (1.544)	-0.377 (1.935)	0.985 (1.374)	0.731 (2.745)	-1.809 (1.291)	-0.201 (1.353)
Enjoy Competition (1-5)	0.065 (1.327)	1.597 (1.304)	2.028 (1.304)	-0.670 (1.386)	1.491 (0.679)**	-0.290 (1.364)
Grit (1-5)	-1.771 (1.857)	-0.723 (2.492)	-1.578 (2.020)	-2.604 (4.957)	-3.651 (1.481)**	2.727 (2.189)
Belief on rel. performance (0-1)	23.558 (4.477)***	24.243 (4.625)***	22.582 (4.339)***	26.251 (8.139)***	19.974 (3.870)***	18.966 (3.573)***
Patience Index	1.188 (1.094)	0.548 (1.227)	-0.498 (1.473)	-2.990 (2.193)	0.018 (0.753)	0.603 (1.233)
Risk Taking Index	1.522 (1.240)	-0.994 (1.213)	-1.642 (1.619)	-0.284 (2.770)	0.208 (0.783)	-2.371 (1.194)**
<b>Time Difference between Sessions (in days)</b>						
Time Difference	0.158 (0.102)	-0.068 (0.178)	-0.237 (0.136)*	-0.273 (0.216)	-0.015 (0.099)	0.012 (0.167)
Constant	42.157 (21.330)**	61.891 (26.502)**	55.529 (17.971)***	98.111 (22.943)***	75.193 (14.845)***	79.565 (23.485)***
Num.Obs.	331	327	325	235	458	257
R2 Adj.	0.343	0.340	0.344	0.116	0.472	0.472

Note: Table shows OLS regressions on performance in part 2 split by treatment and payment. Standard errors (in parentheses) clustered on the session level.

Table A12: Productivity by treatment and incentive scheme - Robustness towards time difference between sessions.

	Exogenous			Endogenous		
	Payment			Payment		
	Fixed	Variable	Tournament	Fixed	Variable	Tournament
(I)	(II)	(III)	(IV)	(V)	(VI)	
<b>Skills</b>						
Grade German	-0.430 (0.980)	-0.329 (1.763)	0.246 (1.174)	-2.894 (1.344)**	0.003 (0.904)	2.976 (1.070)**
Grade Math	4.317 (0.977)***	2.219 (0.821)***	1.297 (0.979)	1.660 (1.483)	1.900 (0.729)***	0.428 (1.002)
IQ (Raven 0-10)	0.211 (0.550)	0.432 (0.646)	1.089 (0.642)*	1.238 (1.417)	0.935 (0.516)*	1.573 (0.806)*
Productivity (resid.)	1.722 (0.220)***	1.576 (0.267)***	1.923 (0.307)***	1.153 (0.582)**	2.439 (0.202)***	2.204 (0.281)***
<b>Demographics</b>						
Age rel. to grade mean	-2.495 (0.829)***	-3.612 (1.259)***	-2.403 (1.115)**	0.425 (1.377)	-0.581 (1.114)	-0.007 (1.027)
Female (=1)	4.843 (2.667)*	2.687 (2.478)	1.764 (2.177)	-0.004 (5.355)	4.536 (1.416)***	3.712 (3.017)
Grade (9-13)	4.450 (0.914)***	3.935 (1.596)**	4.815 (1.003)***	2.361 (2.368)	2.711 (0.817)***	4.034 (1.023)***
Number of siblings	-0.271 (1.300)	-1.068 (0.818)	0.723 (0.797)	-1.338 (1.138)	-0.081 (0.709)	0.108 (1.410)
Positive Parenting (0-5)	0.849 (1.352)	-1.076 (1.163)	-2.348 (1.103)**	-2.788 (2.256)	-1.328 (0.557)**	-1.120 (1.154)
Composite SES Index	-0.375 (0.952)	-0.969 (1.006)	-1.364 (1.092)	-0.458 (1.525)	-1.526 (0.557)***	0.891 (1.181)
<b>Personality Traits, Economic Preferences, and Beliefs</b>						
Altruism (0-10)	-0.341 (0.436)	0.757 (0.470)	0.145 (0.567)	0.854 (1.073)	0.117 (0.405)	-0.888 (0.431)**
Extraversion (0-5)	-1.793 (1.256)	-0.127 (1.198)	-0.391 (1.575)	-3.952 (2.143)*	-0.613 (1.312)	1.345 (1.251)
Agreeableness (0-5)	1.404 (2.828)	-1.409 (1.921)	-2.278 (1.849)	-3.180 (4.061)	2.599 (1.444)*	2.689 (1.346)**
Conscientiousness (0-5)	-0.047 (2.395)	0.809 (4.177)	2.157 (2.724)	7.129 (5.280)	2.642 (2.265)	-6.415 (3.033)**
Neuroticism (0-5)	-0.279 (1.220)	-0.509 (1.311)	-1.473 (1.548)	-0.387 (3.543)	0.919 (1.149)	-1.202 (1.849)
Openness (0-5)	1.463 (1.499)	-0.417 (1.999)	0.751 (1.370)	-0.080 (2.643)	-1.712 (1.240)	0.522 (1.277)
Enjoy Competition (1-5)	0.284 (1.400)	1.645 (1.273)	1.775 (1.252)	-0.563 (1.486)	1.711 (0.630)***	-0.234 (1.321)
Grit (1-5)	-2.040 (1.842)	-0.753 (2.528)	-1.996 (1.994)	-2.794 (4.651)	-3.712 (1.563)**	2.013 (1.968)
Belief on rel. performance (0-1)	23.541 (4.605)***	23.920 (4.580)***	23.359 (3.920)***	26.395 (8.220)***	21.504 (3.695)***	19.510 (3.327)***
Patience Index	1.459 (1.227)	0.607 (1.201)	-0.405 (1.525)	-2.226 (2.231)	0.552 (0.638)	0.695 (1.287)
Risk Taking Index	1.758 (1.183)	-0.922 (1.224)	-1.911 (1.643)	-1.077 (2.800)	0.230 (0.808)	-2.591 (1.051)**
<b>Year of Conduction</b>						
Year = 2020	2.722 (4.013)	4.748 (3.140)	9.189 (3.407)***	5.017 (5.647)	2.226 (2.093)	4.806 (3.501)
Year = 2022	-2.357 (2.858)	-2.034 (2.898)	3.017 (2.859)	-8.913 (4.766)*	-8.391 (3.070)***	-10.441 (4.194)**
Constant	40.832 (21.596)*	58.876 (27.026)**	52.230 (17.990)***	84.061 (23.402)***	66.143 (16.091)***	62.678 (16.845)***
Num.Obs.	331	327	325	235	458	257
R2 Adj.	0.342	0.342	0.348	0.128	0.498	0.510

Note: Table shows OLS regressions on performance in part 2 split by treatment and payment. Standard errors (in parentheses) clustered on the session level.

Table A13: Productivity by treatment and incentive scheme - Robustness towards time difference between sessions.

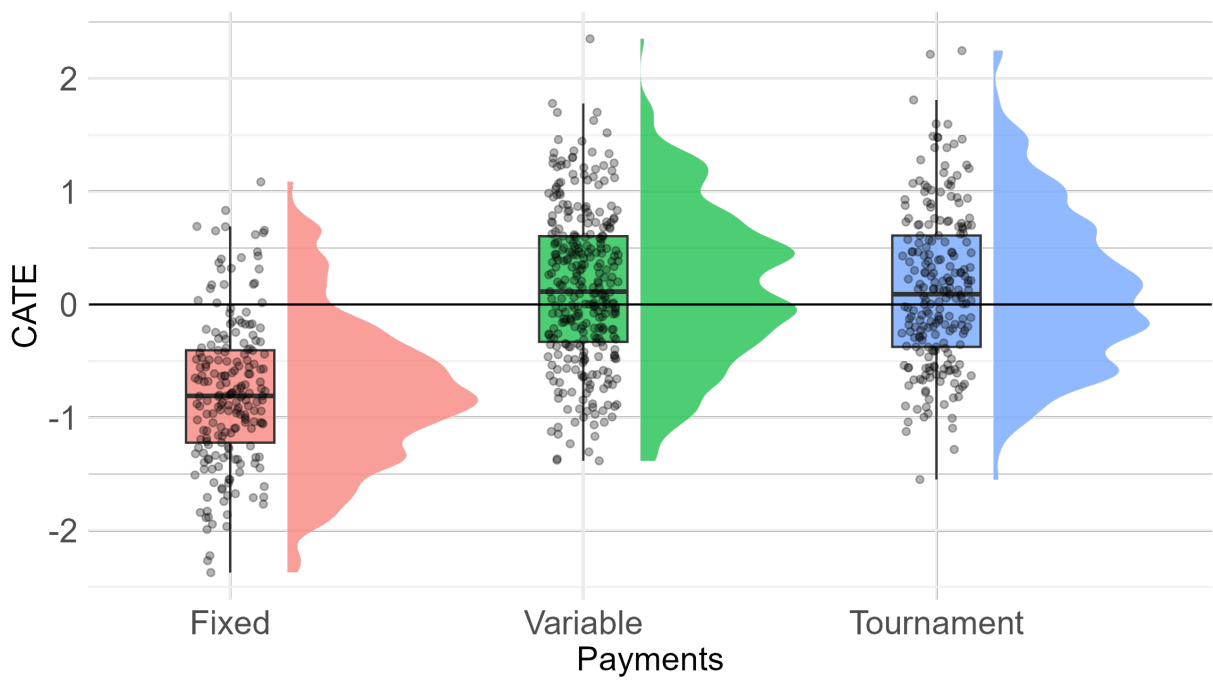


Figure A4: CATE by Payment - Excluding low-performer



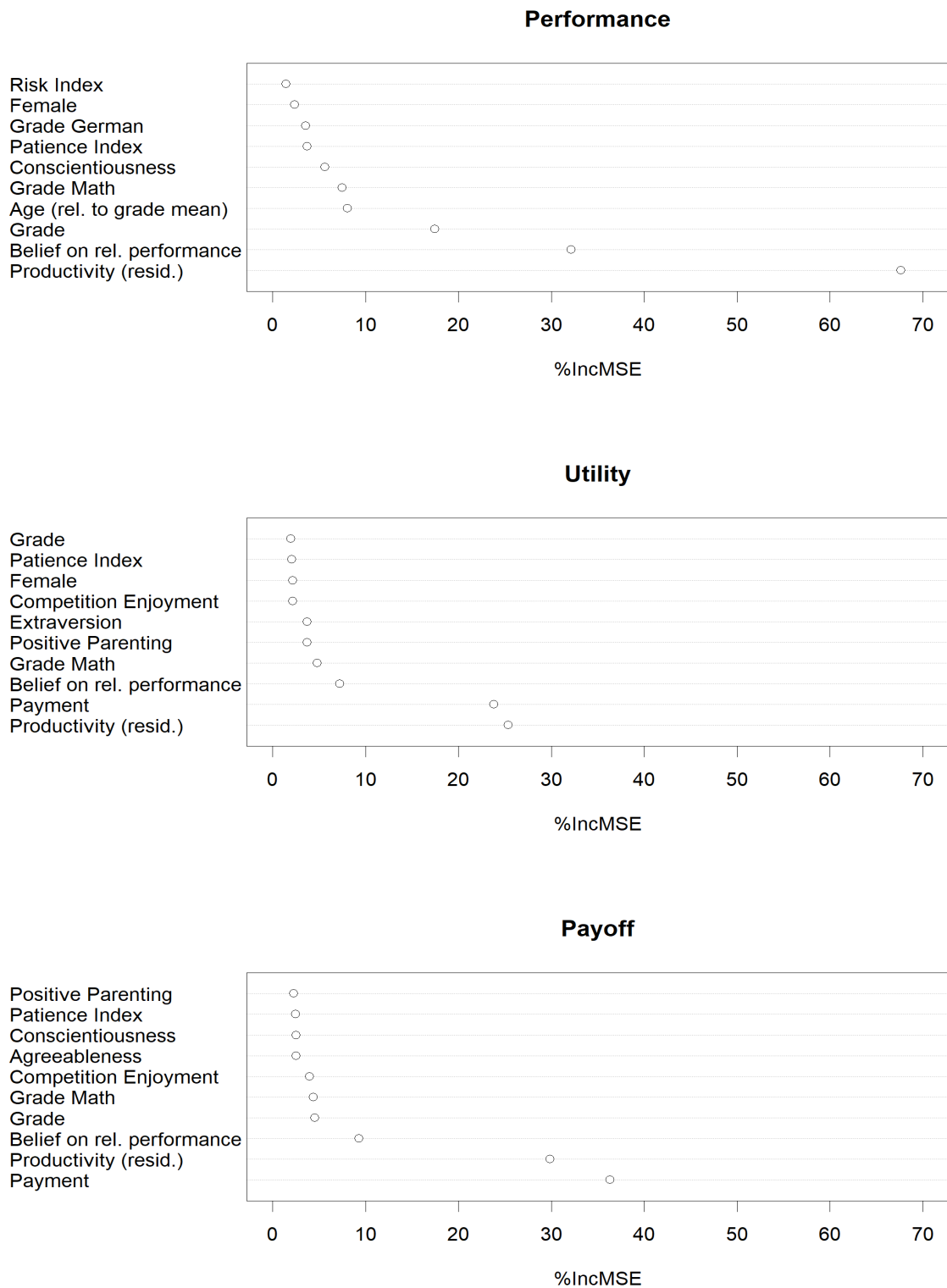


Figure A5: Full Sample - Top 10 variable importance for random forest predictions for optimal assignment of payments. %IncMSE denotes the decrease in accuracy when a given variable is permuted. Higher values translate into a stronger influence in the prediction process. We report results on three prediction outcomes: performance, utility and payoff.

## A.4 List of explanatory variables

Here, we provide a description of the included explanatory variables. All variables were carefully chosen based on their potential to shape earnings and life outcomes, reported in the literature.

**Socio-economic status and other relevant socio-demographic variables.** SES and demographic variables have been shown to be strongly associated with educational outcomes and earnings (Cunha and Heckman, 2007; Heckman, 2006). Besides age and gender, our questionnaire included a number of proxies for SES<sup>21</sup>, which we used to construct three main indexes and one SES component using PCA:

- PISA wealth index: The PISA test provides valuable information to educational researchers and policy makers around the world by comparing countries with regard to their educational system using a variety of educational outcomes. In many ways PISA has emerged as the international benchmark in comparing educational systems (Fuchs and Woessmann, 2008; Hanushek and Woessmann, 2011; Woessmann, 2016). Their SES indicators have often been used for assessing socioeconomic background with teenagers (Hanushek and Woessmann, 2011; West and Woessmann, 2010; Woessmann, 2016). We focus on the family wealth possessions index (WEALTH), which has been validated as a strong and reliable proxy for SES (Schulz, 2006; Rutkowski and Rutkowski, 2013).<sup>22</sup> It includes seven items: *(i)* Do you have a room of your own? Or do you share your room (e.g. with siblings)?; *(ii)* Do you have a link to the Internet at home?; *(iii)* How many cell phones are there at your home?; *(iv)* How many televisions are there at your home?; *(v)* How many computers are there at your home?; *(vi)* How many cars are there at your home?; and *(vii)* How many rooms with a bath or shower are there at your home? In addition to this, we include the number of books available at home, which has been found to alone be another important proxy for socioeconomic status in the PISA test (Woessmann, 2016).
- Family Affluence Scale (FAS) score: This score is also commonly used to elicit SES among school-aged children (Andersen et al., 2008; Boyce et al., 2006; Hartley et al., 2016; Torsheim et al., 2016). The score is similar to the PISA wealth index, and three of the items are the same. It includes four items: *(i)* Do you have a room of your own? Or do you share your room (e.g. with siblings)?; *(ii)* Does your family own a car, van or truck?; *(iii)* How many times did you and your family travel out

---

<sup>21</sup>Given our sample of adolescents, elicited information about household income must be assumed to be very noisy, which is why we use alternatives.

<sup>22</sup>The questions were drawn from PISA tests conducted in 2015. They were accessed from [https://www.oecd-ilibrary.org/education/pisa-2015-assessment-and-analytical-framework/pisa-2015-background-questionnaires\\_9789264255425-8-en](https://www.oecd-ilibrary.org/education/pisa-2015-assessment-and-analytical-framework/pisa-2015-background-questionnaires_9789264255425-8-en).

of Germany abroad for holiday/vacation last year?; and (iv) How many computers does your family own?<sup>23</sup>

- Education and family structure: We follow Kosse et al. (2020) in considering educational and time resources available to the family as important determinants of SES. We classify a participant as low SES if at least one of the following two conditions are fulfilled: (i) neither parents has a college degree; (ii) the participant lives in a single-parent household<sup>24</sup>
- We collected several other relevant variables: number of siblings, zip-code, pocket money, migration background and speaking a different language than German at home (Hansson and Gustafsson, 2013; Woessmann, 2016)
- Socio-economic PCA component: The items of the above listed three socio-economic indexes are used together with the migration indicators to create one component based on principal component analysis. In the principal component analysis, we see eigenvalues of the component in Figure A6. We can see in Table A14 what socio-economic items are binned in the component. When the loading is greater than 0.3, it shows that those items load heavily on the factor.

### Eigenvalues in PCA.

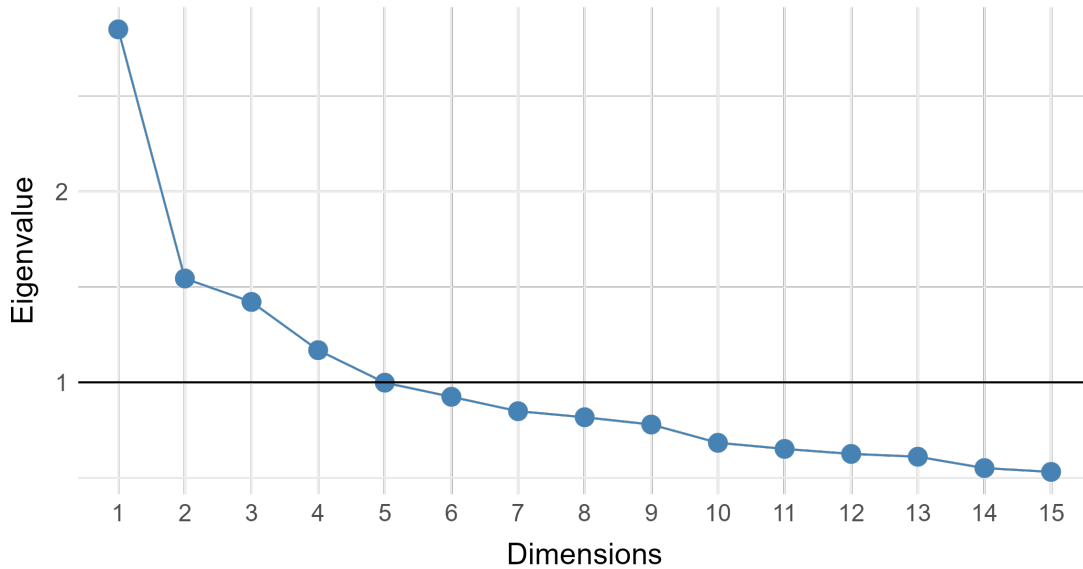


Figure A6: This figure shows the scree plot of eigenvalues of the included fifteen socio-economic variables using principal component analysis.

<sup>23</sup>The composite FAS score is calculated for each adolescent by adding the four items and further categorized into scores below 5, scores between 5 and 7 and scores above 8.

<sup>24</sup>Kosse et al. (2020) consider a third dimension to define SES: household income. While we do not have the actual income, we can use a proxy for household income and compute a similar index.

Socio-economic components	
Number of computers	0.387
Number of cars	0.384
Parents German	0.349
Number of bathrooms	0.305

**Note:** Notes: This table shows the rotated component loading from varimax rotations of principal component analysis of the included fifteen socio-economic variables. Variables with loadings less than 0.3 are excluded from this table.

Table A14: Rotated component loading for socio-economic variables

**Reference level of productivity and stress level.** Baseline measure of performance was captured in Part 1, where a 5 minutes RET paid on a piece-rate was performed. This serves as a proxy for individual’s productivity in playing the real effort task. The reference level of productivity has been found to be important for sorting decisions (Dohmen and Falk, 2011).<sup>25</sup> At the end of the RET (both 5 and 20 minute version) we followed Dohmen and Falk (2011) and elicited self-reported measures of effort, stress, and exhaustion. All three have been found to be higher in pay for performance schemes compared to fixed payment schemes.

**Beliefs.** Overconfidence was computed based on incentivized beliefs. Prior to starting the 20 minutes real effort task in Part 2, we collected information about a participant’s guessed rank in the 5 minute real effort task in Part 1. They got paid according to their guess at the end of the study.<sup>26</sup> Beliefs have been found to be important in sorting decisions, for example in explaining gender differences in sorting into tournaments and differences in productivity. (Bordalo et al., 2019; DellaVigna and Pope, 2017; Dohmen and Falk, 2011; Larkin and Leider, 2012; Reuben et al., 2017).

**Cognitive abilities.** Cognitive ability has been found to be important of school attainment as well as future earnings (Hanushek et al., 2016; Borghans et al., 2008b; Cawley et al., 2001; Segal, 2012). The main proxy for cognitive ability is the score obtained in the Raven’s matrix test administered in Part 1 of the experiment (Raven, 2000). Additionally, we consider self-reported math and German grades. All are expected to be highly correlated with productivity in the real effort task (Dohmen and Falk, 2011).

**Altruism, risk, and time preferences.** Risk and time preferences predict labor market outcomes, educational attainment, income and wealth (Alan and Ertac, 2018; Becker et al., 2012; Bonin et al., 2007; Cadena and Keys, 2015; DellaVigna and Paserman, 2005; Dohmen

<sup>25</sup>Subjects were instructed to solve as many tables as they can, and are given 0.06 cents for each correctly solved table.

<sup>26</sup>See details in the design section.

et al., 2011; Golsteyn et al., 2014; Sutter et al., 2013; Von Gaudecker et al., 2011). They have also been found to be important for different sorting decisions (Bonin et al., 2007; Dohmen and Falk, 2010, 2011). Altruism, risk, and time preferences are measured by using a subset of the global preference survey by Falk et al. (2023, 2018). For both time and risk preferences, we collected multiple measures: a qualitative measure and a quantitative one (staircase). We combined them as proposed by Falk et al. (2018).

**Big five.** Personality traits, such as the big five (Openness, conscientiousness, extraversion, agreeableness and neuroticism), have been shown to be stable traits in affecting performance and life outcomes (Almlund et al., 2011; Akee et al., 2018; Cubel et al., 2016; Deming, 2017; Lindqvist and Vestman, 2011; Segal, 2012). We collected data for all big-five, but mostly focus on conscientiousness and neuroticism as they are found to be consistent predictors of performance in various settings (Borghans et al., 2008a; Donato et al., 2017; Heckman and Rubinstein, 2001).

**Competitiveness.** A large literature documents gender differences in competitiveness. Women avoid competitive schemes, while men are competing too much (Gneezy et al., 2003; Niederle and Vesterlund, 2007, 2010). These gender differences can potentially explain differences in education and labor market outcomes (Gneezy et al., 2003; Niederle and Vesterlund, 2007, 2010). A high level of competitiveness is also a strong predictor for choosing a more prestigious academic track, controlling for ability (Buser et al., 2014, 2017; Reuben et al., 2017) as well as sorting decisions between different payment schemes (Dohmen and Falk, 2011). Competitiveness is measured in our study on the basis of the Competitive Orientation Measure (one single composite scale; see (Newby and Klein, 2014)).

**Parenting style.** Parenting style is important for the academic achievements and future success of children (Doepke et al., 2019; Kosse et al., 2020). We elicited a vital component of parenting style: positive parenting (Essau et al., 2006; Frick, 1991), which indicates the use of positive stimuli and rewards by parents. Recent literature has shown that parental investments have important impact on child cognitive and non cognitive outcomes (Carneiro and Heckman, 2003; Deckers et al., 2017; Doepke et al., 2019).

**Grit.** Grit is defined as perseverance toward a set goal and is seen as being closely related to conscientiousness (Alan et al., 2019). Grit has been found to be predictive of success in a variety of contexts such as through college GPA and educational attainment (Alan et al., 2019; Duckworth et al., 2007). In their study, Alan et al. (2019) finds that students participating in a grit focused intervention chose more challenging tasks and perform better in the real effort task. It was measured in our study by the short-scale Duckworth Grit Index (Dobbie and Fryer Jr, 2015; Duckworth and Quinn, 2009).

## A.5 Instructions

# Choice & Exogenous Treatments

## Instructions for Part 1

### Create your ID

**Experimenter reads aloud:** *[Welcome to the study. This study consists of two sessions: this session today, and another session in which you will participate in the near future. In both sessions you will earn money, please listen carefully to the instructions. For today, you will receive a fixed payment of 4€ if you complete the session. You can also earn additional money depending on your performance in a task that I will explain later. During the session you cannot talk to the other students in this room. This is a very important rule, and if you break it, you will not receive the money that you earned. On the first page, you will be asked to enter your ID as explained on the screen. Your name will never be used during the entire study. Whatever you do, and all the answers you give will only be recorded under your ID. That means that everything you do in the study is going to be anonymous. When you will participate in the second session, you will also be identified via the same anonymous ID code. If you have any questions, please raise your hand and one of us will come to your desk to answer it in private. Please remember that your participation on this study is fully voluntary, and you can decide to quit at any time. If you decide to quit before finishing the study, you are not allowed to leave the room, and you are still required to stay seated at your desk.]*

Welcome to this study. Before we proceed, use the drop-downs to enter your ID. Please double check all your entries before proceeding, as it is very important that your ID is specified correctly.

- Month of birth *[drop down – Jan, Feb, Mar, ...]*
- First and second letter of your mother's first name (or your legal guardian's first name)
- First and second letter of the street where you live

CONFIRM

⇒ ——— Enter ID (first trial) ——— ⇐

## Erstelle deine ID

Willkommen zu dieser Studie! Bevor wir weitermachen, benutze bitte die angezeigten Auswahlmöglichkeiten, um deine ID einzugeben. Bitte überprüfe deine Eingaben anschließend nochmals.

Geburtsmonat:

Erster und zweiter Buchstabe des Vornamens  
deiner Mutter (oder deines (Haupt-)Erziehungsberechtigten):

Erster und zweiter Buchstabe der Straße, in der du wohnst:

**Bestätigen**

⇒ ——— *pop-up* ——— ⇐

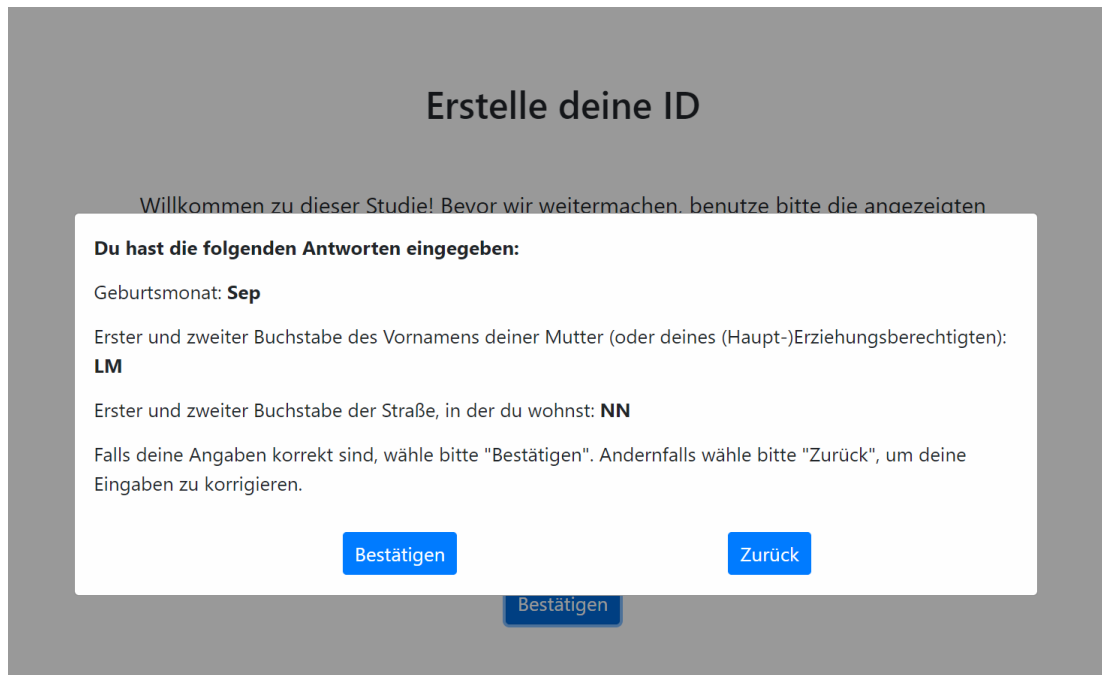
You provided the following answers:

- Month of birth: \_\_\_\_\_
- First and second letter of your mother's first name (or your legal guardian's first name): \_\_\_\_\_
- First and second letter of the street where you live: \_\_\_\_\_

If your answer is correct please press CONFIRM otherwise press BACK to revise your entries.

CONFIRM

BACK



⇒ ——— *wait for all & new screen for double IDs* ——— ⇐

Your ID is the same of someone else in this room. We hence ask you to answer an additional question:

- Last two letter of your first name

CONFIRM



⇒ ——— wait for all & new screen ——— ⇐

## Your Task

### *Experimenter reads aloud:*

You will be shown a set of tables with 1s and 0s, like the one reported below. Your task is to correctly solve as many tables as you can.

Die verbleibende Zeit: 20:00

Anzahl der verbleibenden Versuche: 3

Korrekt gelöste Tabellen: 0

Tabelle 1

0	1	1	0	0	0	0	0
0	0	0	0	1	0	1	1
0	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16  
17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32

Weiter

To correctly solve a table, you have to:

1. **Tap** on all the cells containing a 1, which will highlight them in a dark green color
2. **Count** the correct amount of 1s that you see in the table, and report this amount in the number pad underneath the table.

Be aware, you are **not** allowed to highlight the 0s! If you accidentally highlight a 0, you can tap on the cell again to change it back to grey.

Once you are done with the tapping and you have reported the number, press CONFIRM. You get three tries to solve a table correctly. You will see the amount of remaining tries in the upper-right corner. If you do not manage to solve a table within the three tries, the next table will be shown on your screen. There are no penalties for not solving a table. You can see the amount of correctly solved tables in the upper right corner at any point during the task.

**You have a total of 5 minutes to solve as many tables as you can.** You will be **paid 0.06€ for each table you solved correctly.** For instance, if by the end of the 5 minutes you solve 1 table correctly, you will earn 0.06€. If by the end of the 5 minutes you solve 10 tables correctly, you will earn 10 times 0.06€, so you will earn 0.6€. Or for instance, if by the end of the 5 minutes you solve 100 tables correctly, you will earn 100 times 0.06€, so you will earn 6€.

Before you start with the task you will have one trial round. That means that the first table you solve will not count for money, but will help you get acquainted with the task. After you correctly solve the first table, the 5 minute period will start.

Remember that you are not allowed to talk to the other participants in this room. If you have any question, please rise your hand and one of us will come to your desk to answer it privately. <sup>27</sup>

---

<sup>27</sup>The program is advanced by the experimenter after about 2-3 min (A "continue" button is displayed for the subjects once the experimenter advances the program) and participants are told to click "Continue" once they are ready to continue the experiment.

⇒ ——— wait for all & new screen<sup>28</sup> ——— ⇐

## The real effort task

Die verbleibende Zeit: **5:00** Anzahl der verbleibenden Versuche: **3**  
Korrekt gelöste Tabellen: **0**

**Tabelle 1**

1	0	0	0	0	1	1	0
0	0	0	1	0	0	1	1
0	0	1	0	0	0	1	1
1	1	1	1	0	0	0	1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32

Weiter

<sup>28</sup>A similar table with "Trial round" is displayed. The picture is the same without the remaining time and correctly solved tables. After subjects correctly solve the trial round, they enter a waiting screen which lasts until everybody solves the trial round. Then a new screen appears with a 5 second countdown: "The 5 minute period for solving the task will start in 5, 4, 3... " Following that, Table 1 is displayed and the 5 minute period starts.

⇒ ——— *new screen* ——— ⇐

The following questions are related to the task you completed. Please answer the following questions referring to the task you just solved. Please indicate your answers on a 7 point scale, where 1 means “not at all” and 7 means “very much”:

	1	2	3	4	5	6	7
How much effort did you exert?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How stressed did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How exhausted did you get?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

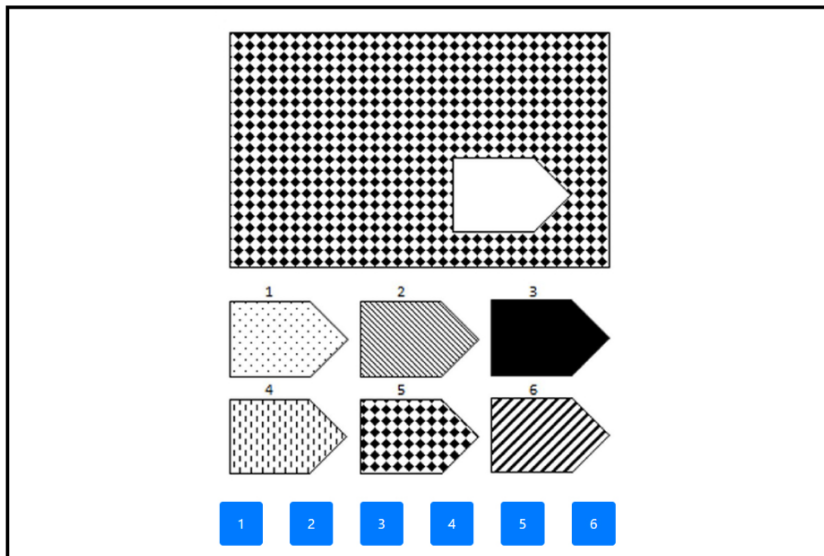
CONTINUE

For the following tasks you have to look at the picture, and find the missing piece of the picture. Once you find it, you need to circle it, as it is shown in the example below. Your goal is to solve as many tasks within 5 minutes as possible. If you have any questions, please raise your hand. If everything is clear, you can start immediately.<sup>29</sup>

## Rätselaufgabe

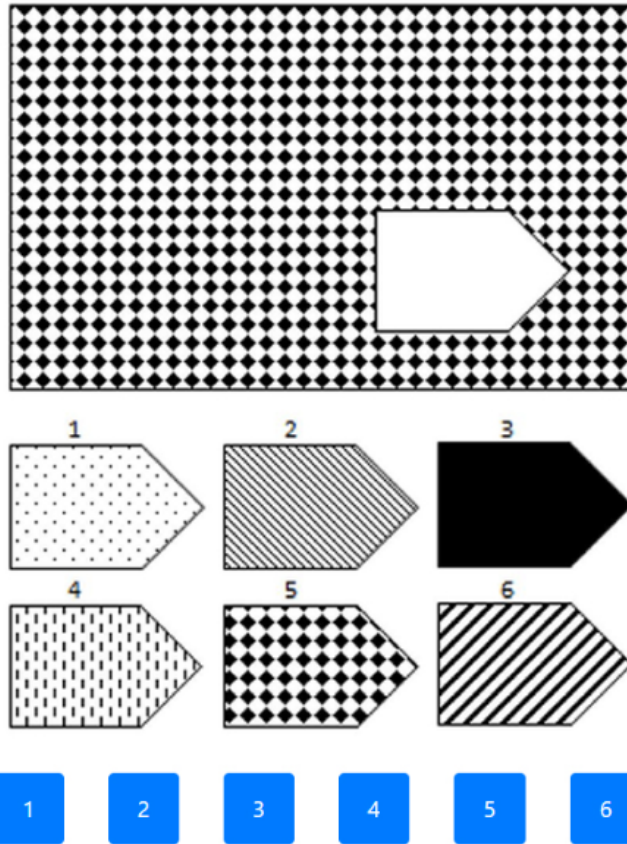
In der nächsten Aufgabe werden Ihnen einige Bilder wie unten abgebildet gezeigt. Sie müssen sich das jeweils gezeigte Bild genau anschauen und das fehlende Teil finden, welches das Bild vervollständigt.

Sobald Sie das fehlende Stück gefunden haben, tippen Sie auf die entsprechende Nummer im unteren Teil des Bildschirms und bestätigen Ihre Auswahl. Ihr Ziel ist es, möglichst viele Bilder innerhalb von 5 Minuten korrekt zu lösen. Sollten Sie eine Frage haben, heben Sie bitte Ihre Hand und ein Mitarbeiter wird zur Klärung zu Ihnen kommen.



NEXT

<sup>29</sup>Instructions are given with an example to make the task clear to the participants. The program is then advanced by the experimenter after they are done with reading the instructions (A "continue" button is displayed for the subjects once the experimenter advances the program) and participants are told to click "Continue" once they are ready to continue with the task.



<sup>30</sup>The matrices are shown and the 5 minutes count-down starts to solve the 10 different matrices. Participants cannot go back once they have submitted an answer for a given task or skip between pictures. If a participant completes the tasks before the 5 minutes, they would have to wait for the other participants to finish.

<sup>31</sup>When subjects have answered all the matrices in the IQ task, they are told to proceed with the remainder of the tasks by themselves.

Suppose you were given the choice between the following: receiving a payment today or a payment in 12 months. We will now present to you five situations. The payment today is the same in each of these situations. The payment in 12 months is different in every situation. For each of these situations we would like to know which you would choose.<sup>32</sup>

## Aufgabe 1

Angenommen, Sie würden vor die Wahl gestellt, eine Zahlung heute oder eine Zahlung in 12 Monaten zu erhalten. Wir nennen Ihnen nun fünf Situationen. Die Zahlung heute ist in allen dieser Situationen identisch. Die Zahlung in 12 Monaten unterscheidet sich in jeder Situation. Für jede dieser Situation würden wir gerne wissen, welche Zahlung sie wählen würden. Bitte gehen Sie davon aus, dass es keine Inflation gibt, somit zukünftige Preise die gleichen sind wie heutige Preise.

Bitte bewerten Sie Folgendes: Würden Sie lieber **100 Euro heute** oder **154 Euro in 12 Monaten** erhalten?

100 Euro heute

154 Euro in 12 Monaten

---

<sup>32</sup>The two different options are shown with the corresponding amounts and for all 5 pages and the different amounts are highlighted to make them salient. Subjects advance to the next page when they click on one of the two buttons. The staircase approach is taken from Falk et al. (2023, 2018).

⇒ ——— *new screen – staircase for risk preferences* ——— ⇐

Please imagine the following situation: You can choose between a sure payment and a lottery. The lottery gives you a 50 percent chance of receiving 300 €. With an equally high chance you receive nothing. Now imagine you had to choose between the lottery and a sure payment. We will present to you five different situations. The lottery is the same in all situations. The sure payment is different in every situation.<sup>33</sup>

## Aufgabe 2

Bitte stellen Sie sich folgende Situation vor: sie könnten wählen zwischen einer sicheren Zahlung eines bestimmten Geldbetrags, ODER einer Verlosung von 300 Euro, bei der Sie die gleichen Chancen hätten, die 300 Euro oder gar nichts zu bekommen. Wir werden Ihnen nun fünf verschiedene Situationen zeigen:

"Was würden Sie bevorzugen?: Würden Sie eine Verlosung mit einer **50-prozentigen Chance 300 Euro** zu bekommen und der gleichen **50-prozentigen Chance nichts** zu bekommen ODER den sicheren Betrag in Höhe von **160 Euro** bevorzugen?"

300 Euro oder 0 Euro

160 Euro sicher

---

<sup>33</sup>The two different options are shown with the corresponding amounts and for all 5 pages and the different amount are highlighted to make them salient. Subjects advance to the next page when they click on one of the two buttons. The staircase approach is taken from Falk et al. (2023, 2018)



⇒ ——— *new screen* ——— ⇐

⇒ ——— **Questionnaire** ——— ⇐

1. Are you female or male?	<input type="radio"/> Female	<input type="radio"/> Male	<input type="radio"/> I don't want to comment
2. What is your zip code?			
3. Where are you born?	<input type="radio"/> In Germany <input type="radio"/> In another EU country <input type="radio"/> In a European country outside of EU <input type="radio"/> In an Asian country <input type="radio"/> In an African country <input type="radio"/> In a South American country <input type="radio"/> In a North American country <input type="radio"/> In Australia		
4. What grade are you in?	<input type="radio"/> Grade 10 <input type="radio"/> Grade 11 <input type="radio"/> Grade 12 <input type="radio"/> Grade 13		
5. Year of birth?			
6. Grade in math? (final grade for last school year)	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> 6		
7. Grade in German? (final grade for last school year)	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> 6		
8. If everything goes as planned, when do you plan to finish the Abitur? (If you don't plan to finish the Abitur, please answer "No plans about finishing the Abitur")	<input type="radio"/> 2019 <input type="radio"/> 2020 <input type="radio"/> 2021 <input type="radio"/> 2022 <input type="radio"/> 2023 <input type="radio"/> No plans about finishing the Abitur		
9. How much pocket money/allowance do you get per week?	0-95 € per week		
10. Do you have a mother/father born outside of Germany?	<input type="radio"/> Both parents born in Germany <input type="radio"/> Mother born outside of Germany <input type="radio"/> Father born outside of Germany <input type="radio"/> Both parents born outside Germany		

<p>11. Do you live together with one or two parents (legal guardians)? (If you live with one parent and his/her partner, please answer: Two parents)</p>	<input type="radio"/> One parent	<input type="radio"/> Two parents	<input type="radio"/> Neither
<p>12. What is the highest education level of your mother?</p>	<input type="radio"/> University or similar <input type="radio"/> High school <input type="radio"/> Middle school or lower <input type="radio"/> No schooling <input type="radio"/> I don't know		
<p>13. What is the highest education level of your father?</p>	<input type="radio"/> University or similar <input type="radio"/> High school <input type="radio"/> Middle school or lower <input type="radio"/> No schooling <input type="radio"/> I don't know		
<p>14. What do you plan to do after you finish high school?</p>	<input type="radio"/> University degree in STEM (Science, Technology, Engineering and Mathematics) <input type="radio"/> University degree outside of STEM <input type="radio"/> Vocational training (Ausbildung) <input type="radio"/> I want to find a job <input type="radio"/> I want to take some time off <input type="radio"/> Voluntary military service <input type="radio"/> I don't know		
<p>15. Do you have any siblings?</p>	<input type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 or more		
<p>16. How many books are there in your home?</p>	<input type="radio"/> 0-10 books <input type="radio"/> 11-25 books <input type="radio"/> 26-100 books <input type="radio"/> 101-200 books <input type="radio"/> 201-500 books <input type="radio"/> More than 500 books		
<p>17. What languages do you speak at home most of the time?</p>	<input type="radio"/> German <input type="radio"/> English <input type="radio"/> Turkish <input type="radio"/> Spanish <input type="radio"/> Italian <input type="radio"/> French <input type="radio"/> Arabic <input type="radio"/> Other		

18. How many times did you and your family travel out of Germany abroad for holiday/vacation last year?	<input type="radio"/> None <input type="radio"/> Once <input type="radio"/> Twice <input type="radio"/> More than twice		
Which of the following are in your home?			
19. A room of your own?	<input type="radio"/> Yes	<input type="radio"/> No	
20. A link to the Internet?	<input type="radio"/> Yes	<input type="radio"/> No	
How many of these are there at your home?			
21. Cell phones?	<input type="radio"/> None <input type="radio"/> One <input type="radio"/> Two <input type="radio"/> Three or more		
22. Televisions?	<input type="radio"/> None <input type="radio"/> One <input type="radio"/> Two <input type="radio"/> Three or more		
23. Computers/PCs?	<input type="radio"/> None <input type="radio"/> One <input type="radio"/> Two <input type="radio"/> Three or more		
24. Cars?	<input type="radio"/> None <input type="radio"/> One <input type="radio"/> Two <input type="radio"/> Three or more		
25. Rooms with a bath or shower?	<input type="radio"/> None <input type="radio"/> One <input type="radio"/> Two <input type="radio"/> Three or more		

Please tell me, in general, how willing or unwilling you are to take risks. Please use a scale from 0 to 10, where 0 means you are "completely unwilling to take risks" and a 10 means you are "very willing to take risks". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

<i>0 = Completely unwilling to take risks</i>	<i>Very willing to take risks = 10</i>
○    ○    ○    ○    ○    ○    ○    ○	○    ○    ○
0    1    2    3    4    5    6    7	8    9    10

We now ask for your willingness to act in a certain way in four different areas. Please again indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing to do so". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

	Completely <b>unwilling</b> to do so	Completely <b>willing</b> to do so
How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?	○ ○ ○    ○    ○ ○ ○ ○ ○ ○    ○	0 1 2 3 4 5 6 7 8 9 10

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please select a number next to each statement to indicate the extent to which you agree or disagree with that statement.

<b>I see Myself as Someone Who...</b>	1. Disagree strongly	2. Disagree a little	3. Neither agree nor disagree	4. Agree a little	5. Agree strongly
1. Is talkative	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Tends to find fault with others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Does a thorough job	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Is depressed, blue	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Is original, comes up with new ideas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Is reserved	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Is helpful and unselfish with others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Can be somewhat careless	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Is relaxed, handles stress well	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Is curious about many different things	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. Is full of energy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. Starts quarrels with others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. Is a reliable worker	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. Can be tense	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Is ingenious, a deep thinker	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Generates a lot of enthusiasm	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. Has a forgiving nature	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. Tends to be disorganized	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. Worries a lot	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. Has an active imagination	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. Tends to be quiet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<b>I see Myself as Someone Who...</b>	1. Disagree strongly	2. Disagree a little	3. Neither agree nor disagree	4. Agree a little	5. Agree strongly
23. Tends to be lazy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. Is emotionally stable, not easily upset	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. Is inventive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26. Has an assertive personality	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
27. Can be cold and aloof	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28. Perseveres until the task is finished	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
29. Can be moody	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30. Values artistic, aesthetic experiences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31. Is sometimes shy, inhibited	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32. Is considerate and kind to almost everyone	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33. Does things efficiently	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34. Remains calm in tense situations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35. Prefers work that is routine	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36. Is outgoing, sociable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37. Is sometimes rude to others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38. Makes plans, & follows through with them	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39. Gets nervous easily	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40. Likes to reflect, play with ideas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41. Has few artistic interests	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
42. Likes to cooperate with others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43. Is easily distracted	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44. Is sophisticated in art, music, or literature	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

The following scale measures aspects of competitiveness. Please read each question carefully and try to answer as honestly as possible. Do not spend too much time on any one item; if trying to decide between two responses, choose the one that first comes to mind.

	1. Strongly disagree	2. Slightly disagree	3. Neither agree nor disagree	4. Slightly agree	5. Strongly agree
1. I like competition.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. I am a competitive individual.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. I enjoy competing against an opponent.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. I don't like competing against other people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. I get satisfaction from competing with others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. I find competitive situations unpleasant.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. I dread competing against other people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. I try to avoid competing with others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. I often try to outperform others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. I try to avoid arguments.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. I will do almost anything to avoid an argument.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. I often remain quiet rather than risk hurting another person.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. I don't enjoy challenging others even when I think they are wrong.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. In general, I will go along with the group rather than create conflict.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

The following are statements about your family. Please rate each item and how often it TYPICALLY occurs in your home.

		1. Never	2. Almost Never	3. Sometimes	4. Often	5. Always
1	Your parents tells you that you are doing a good job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	Your parents reward you or give you something extra to you for behaving well.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	Your parents compliment yuo when you have done something well.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	Your parents prise you for behaving well.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	Your parents hug or kiss you when you done something well.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	Your parents tell you that they like it when you help out around the house.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Please respond to the following 8 items. Be honest – there are no right or wrong answers!

	1. Not like me at all	2. Not much like me	3. Somewhat like me	4. Mostly like me	5. Very much like me
1. New ideas and projects sometimes distract me from previous ones.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Setbacks don't discourage me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. I have been obsessed with a certain idea or project for a short time but later lost interest.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. I am a hard worker.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. I often set a goal but later choose to pursue a different one.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. I have difficulty maintaining my focus on projects that take more than a few months to complete.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. I finish whatever I begin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. I am diligent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

⇒ ——— *new screen at the end of part 1* ——— ⇐

Thanks for taking part in the study.

In the first task you solved \_\_\_\_\_ tables correctly.

You earnings for this task are: \_\_\_\_\_ € (rounded up at the 10 cents)

In addition, you earned a 4€ fee for taking part in the study.

Your total earnings for today are: \_\_\_\_\_ €

Please remain seated and remember that you are not allowed to talk to the other participants.  
One of the experimenters will come to your desk to give you your earnings.

# Choice Treatment<sup>34</sup>

## Instructions for Part 2<sup>35</sup>

### Create your ID

***Experimenter reads aloud:** [In this study you will earn money, so please listen carefully to the instructions. During the study you cannot talk to the other students in this room. This is a very important rule, and if you break it, you will not receive the money that you earned during the study. On the first page, you will be asked to enter your ID as explained on the screen. Your name will never be used during the study. Whatever you do, and all the answers you give will only be recorded under your ID. That means that everything you do in the study is going to be anonymous! If you have any questions, please raise your hand and one of us will come to your desk to answer it in private. Please remember that your participation on this study is fully voluntary, and you can decide to quit at any time. If you decide to quit before finishing the study, you are not allowed to leave the room, and you are still required to stay seated at your desk.]*

Welcome to this study! Before we proceed, use the drop-downs to enter your ID. Please double check all your entries before proceeding, as it is very important that your ID is specified correctly.

- Month of birth [drop down – Jan, Feb, Mar, ...]
- First and second letter of your mother's first name (or your legal guardian's first name)
- First and second letter of the street where you live
- Last two letter of your first name <sup>36</sup>

CONFIRM

---

<sup>34</sup>Same instructions as for exogenous, except the subjects face no choice screen, and only information about the relevant payment scheme is displayed.

<sup>35</sup>The original German instructions and available upon request from the authors.

<sup>36</sup>Extra question in case of double ID.

⇒ ——— Enter ID (first trial) ——— ⇐

## Erstelle deine ID

Willkommen zu dieser Studie! Bevor wir weitermachen, benutze bitte die angezeigten Auswahlmöglichkeiten, um deine ID einzugeben. Bitte überprüfe deine Eingaben anschließend nochmals.

Geburtsmonat:	<input type="text" value="Sep"/>
Erster und zweiter Buchstabe des Vornamens deiner Mutter (oder deines (Haupt-)Erziehungsberechtigten):	<input type="text" value="L"/> <input type="text" value="M"/>
Erster und zweiter Buchstabe der Straße, in der du wohnst:	<input type="text" value="N"/> <input type="text" value="N"/>

**Bestätigen**

⇒ ——— *pop-up* ——— ⇐

You provided the following answers:

- Month of birth: \_\_\_\_\_
- First and second letter of your mother's first name (or your legal guardian's first name):  
\_\_\_\_\_
- First and second letter of the street where you live: \_\_\_\_\_

If your answer is correct please press CONFIRM otherwise press BACK to revise your entries.

CONFIRM

BACK

**Erstelle deine ID**

Willkommen zu dieser Studie! Bevor wir weitermachen, benutze bitte die angezeigten

**Du hast die folgenden Antworten eingegeben:**

Geburtsmonat: **Sep**

Erster und zweiter Buchstabe des Vornamens deiner Mutter (oder deines (Haupt-)Erziehungsberechtigten):  
**LM**

Erster und zweiter Buchstabe der Straße, in der du wohnst: **NN**

Falls deine Angaben korrekt sind, wähle bitte "Bestätigen". Andernfalls wähle bitte "Zurück", um deine Eingaben zu korrigieren.

⇒ ——— new screen ——— ⇐

⇒ ——— wait for all & new screen ——— ⇐

## Your Task

**Experimenter reads aloud:** [I will now explain you the task in which you can earn money. Some of you have already seen the task as you did it the first time we came to your class. But some of you were not here; to be certain that you all know the task, I will explain it in detail again. Please follow the instructions carefully.]

You will be shown a set of tables with 1s and 0s, like the one reported below. Your task is to correctly solve as many tables as you can.

The screenshot shows a task interface with the following elements:

- Top left: "Die verbleibende Zeit:" with a timer showing "20:00".
- Top right: "Anzahl der verbleibenden Versuche: 3" and "Korrekt gelöste Tabellen: 0".
- Center: "Tabelle 1" with a 4x8 grid of cells. The grid contains 0s and 1s. The 1s are located at (row, column) positions: (1,2), (1,3), (2,5), (2,7), (2,8), (3,2), (3,3), (3,4), (3,5).
- Bottom: A number pad with digits 1-32 and a blue "Weiter" button.

0	1	1	0	0	0	0	0
0	0	0	0	1	0	1	1
0	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0

Number pad: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32

Buttons: Weiter

To correctly solve a table, you have to:

1. **Tap** on all the cells containing a 1, which will highlight them in a dark green color;
2. **Count** the correct amount of 1s that you see in the table, and report this amount in the number pad underneath the table.

Be aware, you are **not** allowed to highlight the 0s! If you accidentally highlight a 0, you can tap on the cell again to change it back to grey.

Once you are done with the tapping and you have reported the number, press CONFIRM. You get three tries to solve a table correctly. You will see the amount of remaining tries in the upper-right corner. If you do not manage to solve a table within the three tries, the next table will be shown on your screen. There are no penalties for not solving a table. You can see the amount of correctly solved tables in the upper right corner at any point during the task. **You have a total of 20 minutes for solving the task.**

Remember that you are not allowed to talk to the other participants in this room. If you have any question, please raise you hand and one of us will come to your desk to answer it privately.<sup>37</sup>

---

<sup>37</sup>The program is advanced by the experimenter after about 2-3 min (A "continue" button is displayed for the subjects once the experimenter advances the program) and participants are told to click "Continue" once they are ready to continue the experiment

## Guessing Task

Before explaining how you will be paid for the task, you have a chance to earn some additional money.

In this room, there are XY students (including you) that were present also during our previous visit. You all performed the task for 5 minutes the other time.

In the 5-minute version of the task, **you correctly solved X tables.**

We ranked you and the other participants present in the previous visit.<sup>39</sup> You were all ranked based on the number of tables correctly solved in 5 minutes. For example, position number 1 is for the one who solved the most tables, position number 2 is for the one who solved the second most tables, and so on, with the last position XY for the one who solved the least.

**We would now like you to guess your position in the ranking.**

**If you were to guess the correct number, you earn 2€. If you come within up to 5 positions (higher or lower), you will earn 0.50€ that will be added to your total earnings for today's session.**

- ————— +  
1 ||| ||| ||| ||| ||| ||| ||| ||| XY

**I think I ranked number .....**

CONFIRM<sup>40</sup>

---

<sup>38</sup>Screen only appears for subjects that are present in both sessions.

<sup>39</sup>Participants that are present in both sessions are ranked by standard competition ranking.

<sup>40</sup>Participants need to touch the slider to activate it. They can adjust the number either by touching the slider or clicking on the + and - signs at the ends of the slider. Absolute numbers of the different options for the ranking appear after the participant click on the slider. "I think I ranked number ...." only appears when participants click on the slider with the number of the ranking clicked on.



## Your Earnings<sup>41</sup>

You can determine the payment mode yourself. In particular, you can choose between three alternative payment modes.

**Fixed Payment.** When the 20 minutes are up, you will receive 6.5€, independent of the number of tables you solved correctly.

**Variable Payment.** When the 20 minutes are up, you will be paid 0.06€ for each table you solved correctly.

**Tournament.** When the 20 minutes are up, you will be paid either 0.08€ or 0.04€ *for each table* you solved correctly. To establish whether you will be paid 0.08€ or 0.04€ per correct table, your performance will be compared with one other student in this room, whose payment will also be determined in the same way. At the end of the 20 minutes, if you solved more tables than this other student matched with you, you will get 0.08€ per correct table. If instead you solved less tables than this other participant matched with you, you will get 0.04€ per correct table. If you and this other participant solved the same number of tables, the computer will randomly determine if you are paid 0.08€ or 0.04€ per correct table.

For your information, in the first visit you have solved \_\_\_\_\_ in \_\_\_\_\_ minutes. <sup>42</sup>

Before choosing your payment mode, please answer a few control questions designed to make sure you understood how the earnings are computed. If you have any questions, please raise your hand and wait for an experimenter to come to your desk. <sup>43</sup>

1. In the **fixed payment** alternative, if you solve 10 tables correctly by the end of the 20 minutes, how many Euros will you get?
  - a I will get 0.60€
  - b I will get 6.50€**
  - c I will get 80.00€
2. In the **fixed payment** alternative, if you solve 1000 tables correctly by the end of the 20 minutes, how many Euros will you get?
  - a I will get 0.60€

---

<sup>41</sup>In the exogenous treatment, subjects would only be able to see the paragraph explaining the payment scheme that they were assigned, and would only receive control questions referring to that payment scheme.

<sup>42</sup>This info was displayed only if the ID was present in first study and it is unique in second study.

<sup>43</sup>The correct answers are marked here in bold for display.

- b I will get 6.50€
  - c I will get 80.00€
- 3. In the **variable payment** alternative, if you solve 10 tables correctly by the end of the 20 minutes, how many Euros will you get?
  - a I will get 0.60€
  - b I will get 6.50€
  - c I will get 90.00€
- 4. In the **variable payment** alternative, if you solve 1000 tables correctly by the end of 20 minutes, how many Euros will you get?
  - a I will get 0.60€
  - b I will get 6.50€
  - c I will get 60.00€
- 5. In the **tournament payment** alternative, if you solve 1000 tables correctly, and the student matched with you solves 10 tables correctly by the end of the 20 minutes, how many Euros will you get?
  - a I will get 0.40€
  - b I will get 6.50€
  - c I will get 80.00€
- 6. In the **tournament payment** alternative, if you solve 10 tables correctly, and the student matched with you solves 1000 tables correctly by the end of the 20 minutes, how many Euros will you get?
  - a I will get 0.40€
  - b I will get 6.50€
  - c I will get 80.00€

⇒ ——— new screen ——— ⇐

## Choice of Payment Mode

Bitte wählen Sie eine Auszahlungsvariante und drücken Sie auf weiter.		
Feste Auszahlung	Variable Auszahlung	Wettbewerb
Sobald die 20 Minuten abgelaufen sind, erhalten sie 8€. Der Betrag ist dabei unabhängig von der Anzahl der von ihnen gelösten Tabellen.	Sobald die 20 Minuten abgelaufen sind, erhalten sie 0,06€ für jede korrekt gelöste Tabelle.	Sobald die 20 Minuten abgelaufen sind, erhalten sie entweder 0,1€ oder 0,04€ für jede korrekt gelöste Tabelle. Um zu bestimmen, ob ihre Auszahlung 0,1€ oder 0,04€ für jede gelöste Tabelle beträgt, wird ihre Leistung mit einer/m anderen Person in diesem Raum verglichen, der/die auch die Wettbewerbsauszahlung ausgewählt hat. Sollten sie nach Ablauf der 20 Minuten mehr Tabellen als diese andere Person korrekt gelöst haben, werden ihnen 0,1€ für jede korrekt gelöste Tabelle ausgezahlt. Falls sie allerdings weniger Tabellen als diese

⇒ ——— wait for all & new screen ——— ⇐

### Show if only one person chose tournament

You are the only one who chose tournament. Unfortunately, it is not possible to match you with another student in this room. Please choose again, this time between fixed and variable payment.<sup>44</sup>

⇒ ——— wait for all & new screen ——— ⇐

Countdown. The task will start in 10, 9, 8...seconds.

<sup>44</sup>A menu with the two possible choices are shown to the participant.

⇒ ——— new screen – the real effort task ——— ⇐

Die verbleibende Zeit: **20:00** Anzahl der verbleibenden Versuche: **3**  
Korrekt gelöste Tabellen: **0**

### Tabelle 1

1	0	0	0	0	1	1	0
0	0	0	1	0	0	1	1
0	0	1	0	0	0	1	1
1	1	1	1	0	0	0	1

1234567891011121314151617181920212223242526272829303132

Weiter

⇒ ——— new screen ——— ⇐

Before proceeding to the final payments, please answer the following questions referring to the task you just solved. Please indicate your answers on a 7 point scale, where 1 means “not at all” and 7 means “very much”:

	1	2	3	4	5	6	7
How much effort did you exert?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How stressed did you feel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How exhausted did you get?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

NEXT

⇒ ——— *new screen at the end of the task* ——— ⇐

Thank you for taking part in our study. The task is now over.

You solved \_\_\_\_\_ tables correctly.

Your earnings for this task are: \_\_\_\_\_ €

In addition, you earned a 1€ fee for taking part in the study.

In the guessing task you earned ..... €

**[Only for Tournament.]** You solved more/less/the same number of tables than/as the student you are compared to.

**[Only in case of tie.]** The computer randomly determined that you earn 0.08/0.04€ per solved table.

Your earnings for the task are: \_\_\_\_\_ €  
(rounded up at the 10 cents)

Please remain seated and remember that you are not allowed to talk to the other participants. One of the experimenters will come to your desk to give you your earnings.