universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

"Effects of Information on Competitive Behavior and Decision Making"

verfasst von / submitted by

Dominik Stelzeneder, Bakk. MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 794 370 140
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Volkswirtschaftslehre
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Martin Kocher Univ.-Prof. Dipl.-Math. Karl Schlag, PhD

Acknowledgments

I am grateful to my supervisors Martin Kocher and Karl Schlag for supporting me throughout my PhD studies. This thesis could have not been written without your constant support, your trust and your encouragement. It was an honor and a privilege to work with you and to learn from you.

I also thank the faculty of the University of Vienna for the lively discussions and for the interest in my research. A particular debt of gratitude is owed to Philipp Schmidt-Dengler for his support and guidance.

It was a great pleasure to work with three wonderful coauthors on two of my chapters. Thank you, Boris, Kai and Seb! I would also like to thank Markus Popp, Helwin Prohaska, Katharina Weissenböck, apprentices at HBV, and principals of vocational schools for their support.

Furthermore, I am thankful to all my fantastic colleagues at VGSE. Doing research and writing it up is a much more pleasant experience when the ups and downs can be shared with somebody. They were always available for a cup of coffee and a pleasant break from research. Many of them have become good friends.

Finally, I would like to thank my family and friends for all the fun and the good time we have had together. And I thank Sophie and David for everything.

Contents

Introduction

Experimental economics is a vibrant and versatile field/method in economics. More than 200 experimental papers were published only in the five leading economic journals from 2010 to 2021 (Congiu, Nunnari, and Sandri 2022). Moreover, these papers were highly influential, as the high average number of citations compared to papers in other fields suggests (Fréchette, Sarnoff, and Yariv 2022). The success of experimental economics goes hand in hand with the breadth of research topics that experimental papers cover. Gender economics, political economy or health and educational economics are just a few of the areas that experimentalists have contributed to.

This dissertation consists of three independent chapters in the field of experimental economics. The range of research topics that this doctoral thesis covers is typified by the wide variety of topics that experimental economists work on. In the first chapter of this thesis, I study the direct causal effects of schooling on political attitudes of vocational students in Austria. In the second chapter, which is joint work with Kai Barron and Sebastian Schweighofer-Kodritsch, I study the effects of frames on men's and women's competitiveness. In the third chapter, which is joint work with Boris Knapp, I study how two informational policies mitigate the negative effects of misinformation on purchasing decisions in consumer markets. All three chapters – as diverse as the research topics therein are – are connected by the experimental methods that I apply to contribute to the economics literature.

The first chapter of my dissertation contributes to the literature at the intersection of education economics and political science. A large body of research documents a strong relationship between citizens' educational attainment and their political attitudes. Education positively correlates with political interest, participation in elections and support for the political left. Based on a quasi-experiment, I study whether schooling causally and directly affects political attitudes of vocational students in Austria.

In this chapter I exploit that vocational students in Austria are as good as randomly assigned to different classes and school terms, respectively. Vocational students are both employees of their training company and they attend vocational school during ten weeks in each year of training. In school, they receive general education and practical training

that depends on their vocation. In a field survey in 2019, I collected data on political attitudes of two groups of vocational students. The first group was surveyed at the end of the first school term and the other group was surveyed just a few days later, at the beginning of the second school term. The first group of vocational students had thus been in school for ten weeks whereas the other group was at work during that time.

I observe that vocational students who were in school were significantly more interested in politics compared to their counterparts who were at work instead. Furthermore, I can show that vocational students who were in school when a major political scandal in Austria was uncovered, reported significantly different party preferences even five months after that event. In spring 2019 the so called *Ibiza affair* was uncovered, which eventually caused the collapse of the government coalition and led to snap elections. Vocational students who went to school when the affair was uncovered were significantly more likely to support no party at all and significantly less likely to support the *Austrian Freedom Party* (FPÖ) that caused the affair. This finding suggests that the general political situation when school is attended matters.

The second contribution of my dissertation lies in the area of gender economics. It is well known that labor market outcomes still differ between men and women. Women earn less than men on average and their chances to become members of corporate boards are lower, for instance (Ciminelli, Schwellnus, and Stadler 2021 and Deloitte 2022). Potential explanatory factors that have gained particular attention in the recent literature are gender differences in competitiveness. Gneezy, Niederle, and Rustichini (2003) showed in a laboratory experiment that men performed better in a competition than women, even though their performance did not differ in a noncompetitive setting. In addition, Niederle and Vesterlund (2007) showed that men are more likely to enter a competitive environment, given a noncompetitive alternative. Factors like risk aversion, performance differences and self-confidence could not entirely account for the higher tournament entry rate of men.

Together with my coauthors Kai Barron and Sebastian Schweighofer-Kodritsch, I study whether the perception and, in particular, the gender stereotype of a task affects men's and women's competitiveness. If the stereotype matters, then dismantling gender stereotypes will likely decrease gender differences in competitiveness and thus also improve women's labor market outcomes. In our experiment we vary the framing of the task that we use in order to study the effect of how the perception of a task affects competitiveness. We use a simple pattern matching task that we refer to as either a *creativity task* or an *intelligence task*.

We find that both men and women performed significantly better in the competition when the task was framed as an *intelligence* task and not as a *creativity* task. Further-

more, the perceived gender stereotype of the task was affected by our treatment. Interestingly, there was no evidence for *differences* in framing effects on men's and women's competitiveness. Men's and women's performance did not differ in the tournament, and decisions to either enter a competition or to select a non-competitive alternative were not affected by our treatment. Our findings suggest that the gender stereotype of a task does not affect the competitiveness of men or women. Measures that help to dismantle domain specific gender stereotypes will thus likely not be effective in decreasing the gender gap in competitiveness and consequent differences in labor market outcomes.

In the third chapter of my thesis, my coauthor Boris Knapp and I contribute to the literature on decision-making under uncertainty. Suppose a consumer has to make a purchasing decision based on a potentially faked review. Making the right decision is a challenge that we conceptualize as a two-step problem. First, the consumer has to form a belief about the trustworthiness of the review. Second, she has to update her prior belief accordingly. Errors in each of these steps might lead to a suboptimal decision. Understanding how these two channels contribute to suboptimal decisions has direct implications for consumer protection policies.

We conducted an online experiment to study the extent to which each of the two channels drives a consumer's decision error. Our experiment is based on a simple signaling game in which a sender sends a costless signal about the quality of a good to a receiver. After observing the signal, the receiver decides either to purchase the good or to take his outside option. Two features of our experiment allow us to assess the extent to which the receiver deviates from her individually optimally decision. First, we elicit the receiver's valuation of the good and not only her decision whether she purchases it or not given a fixed outside option. Second, we elicit her actual valuation of the good, i.e. her valuation if she would have (i) had the correct belief about the trustworthiness of the signal and if she would have (ii) been perfectly able to update her belief accordingly.

We find that inability to correctly compute posterior probabilities significantly impairs decisions and that this effect is heterogeneous across gender. Incorrect beliefs about the trustworthiness of the signal did not increase receivers' decision errors further.

Chapter 1

Does Schooling Affect Political Attitudes? Quasi-Experimental Evidence[1]

In the first chapter of my dissertation I study the direct causal effects of schooling on political attitudes of vocational students in Austria. I exploit that classes of apprentices of the same grade level and vocation are as good as randomly assigned to different school terms. This allows to compare apprentices who were at school for ten weeks with apprentices who were at work in their training firms during that time. I find that schooling has a positive direct causal effect on political interest of vocational students. This increase in political interest is, however, not accompanied by a significant increase in voting intention. Furthermore, my results suggest that apprentices who went to school while being exposed to a political affair support different parties than those apprentices who were exposed to the affair at work.

1.1 Introduction

A large scientific literature about the effects of education on political preferences, voter turnout, and political interest demonstrates how important schools are for political opinion formation and for a lively democracy in general.[2] It was shown that education increases people's political interest and it is a powerful predictor for participation in elections (Milligan, Moretti, and Oreopoulos 2004, Sondheimer and Green 2010, Smets and

[1] A previous version of this chapter was published as Stelzeneder (2023) in the Vienna Economics Papers Working Paper Series.

[2] See for example Hyman (1979), Nie, Junn, and Stehlik-Barry (1996), Weakliem (2002), Berinsky and Lenz (2010), Gerber et al. (2010), Dunn (2011), Persson (2014), Friedman et al. (2015), Persson, Lindgren, and Oskarsson (2016), Croke et al. (2016), Larreguy and Marshall (2017), Solis (2017), Dang (2018).

Ham 2013). Furthermore, education seems to causally steer people to the political right, contrary to what correlational evidence would suggest (Marshall 2016, Meyer 2017 and Marshall 2019).

However, a direct link between schooling and changes in students' political attitudes has not been found yet. The studies cited above do not distinguish between direct effects of schooling and the indirect effects, they solely report the total effects of schooling. This means that the exact mechanism remains unclear since it could be just other, indirect (i.e. mediating) factors that cause the effects that were found. Obtaining a higher level of educational attainment increases income (Mincer 1974), it affects who people spend time with (Rockwell 1976, Chadwick and Solon 2002) and it can even affect health status in the long run (Meghir, Palme, and Simeonova 2018). As a consequence of these changes, political attitudes could be affected as well. In order to understand whether and how schooling per se shapes students' political attitudes, identifying the direct effects seems relevant.

In this paper I use a quasi-experimental design to study whether and how schooling directly affects political attitudes. My analysis is based on data from a field survey among vocational students in Austria. In the Austrian dual education system, vocational students are in a training and work relationship with their training company and they are students in a vocational school. They receive practical training in their company and ten weeks of general education as well as additional theoretical education and practical training in school during each year of training.

I exploit that vocational students in Austria are as good as randomly assigned to different classes and school terms, respectively. This allows comparing the political attitudes of vocational students who only differ in their period of schooling but not with respect to grade level, profession, expectation about the future or other characteristics. This means that it is indeed only the direct effect of schooling that is identified and not any other, indirect effect.

The potential drivers for direct effects of schooling are manifold. It might be the peers in school who affect political attitudes of a student. Discussions of political issues might lead to higher political interest and to more knowledge about politics. Or it might be the teachers who pass on their passion for democracy to students. Attending the school subject *civic education* might be another potential factor that could expand students' political knowledge and trigger them to participate in elections. But many other school specific factors are plausible drivers for changes in political attitudes of students as well.

Isolating the direct effects of schooling from the indirect effects is trickier than one might think (Green, Ha, and Bullock 2010, Clarke 2005, Gillen, Snowberg, and Yarif 2019). Attempts to divide total effects econometrically into a direct and an indirect part

are prone to bias. Omitted variables are one particular issue in this context because schooling has effects on so many different variables that in turn could all affect political attitudes. "Controlling" for these variables in a regression does not seem feasible. It is hard to imagine first, that we could know about all the relevant variables and second, that we are able to measure them without measurement error. Both conditions would be necessary to get unbiased results. But there is also another, more fundamental issue when political attitudes of people who went to school for a different period of time are compared. Even in the absence of any self-selection issue or the factual presence of any of the indirect factors that were mentioned above, future courses of life will differ between people under comparison. This could lead to different expectations about future prospects that might already cause political attitudes to change. Standard econometric approaches reach their limits under these circumstances. Comparing the political attitudes of two groups of students who only differ in their period of schooling bot not with respect to other characteristics is a viable possibility to overcome these issues.

In a field survey in school year 2019, I collected data on political attitudes and on numerous background characteristics of two groups of vocational students. One group was surveyed at the end of the first school term and the other at the very beginning of the second term that started just a few days later. Each term lasts for ten weeks, so at the end of the first school term one group of vocational students had been at school for ten weeks, whereas the other had been at work during that time.

I observe a significant positive effect of ten weeks of schooling on political interest, moderate positive effects on students' likelihood to support the conservative party (ÖVP) and no significant effect on the intention to participate in an election. Furthermore, I can show that the general political situation when school is attended matters. Vocational students who were at school when a major political scandal in Austria was uncovered (the so called *Ibiza affair* in spring 2019) reported significantly different party preferences even five months after that event, compared to vocational students who went to work during that time. They were significantly more likely to support no party at all and significantly less likely to support the *Austrian Freedom Party* (FPÖ) that caused the affair.

What could be the mechanism behind the direct effects of schooling on political attitudes? Suggestive evidence points towards the role of teachers as the driver for the positive effect on students' political interest. When apprentices attend school, they talk more often with teachers and with fellow students about politics and less often with workmates. Among these three groups, apprentices perceive teachers as most interested in politics (followed by workmates and fellow students). Furthermore, spending more time with peers does not significantly affect a student's political interest. I, therefore, conjecture that it is the teachers who are able to pass on part of their political interest to their

students. Results for how students' party support could have been affected by schooling are ambiguous. Only weak indications suggest that attending the school subject *civic education* affects an apprentice's likelihood to support the ÖVP. I do not find evidence for an effect of schooling on knowledge about the positions of Austrian's political parties. Also the quality of vote, measured by the match of the political opinions of an apprentice and the positions of the party he or she supports, does not improve significantly.

My paper makes two contributions to the literature studying the causal effects of schooling. First, I add to the literature by showing that schooling has a direct and positive effect on political interest of vocational students. Other studies investigate the total effects of schooling, which does not allow for differentiating between the direct and indirect effects. Two papers in the literature are closely related to this contribution. Persson, Lindgren, and Oskarsson (2016) exploit the exogenous variation related to school entry age to estimate the effects of one additional year in school on civic values and on the intention to participate in an election. In contrast to my paper they do not study the effect of schooling on *political interest* or on *support for a political party*. Furthermore, their treatment and control groups slightly differ since they compare students who are among the oldest in a class with students who are among the youngest in another class. In the second paper, Solis (2017) does not investigate the effect of schooling but of higher education on students. He exploits the existence of a discontinuity in the eligibility criteria for college loans in Chile to investigate the effect of college education on political participation. In line with my results, neither of the two studies finds a causal effect of education on political participation.

Second, my findings suggest that the general political situation under which schooling takes place matters. All apprentices in my study were exposed to the so-called *Ibiza affair*. But it made a large difference whether an apprentice was going to school or to work while being exposed to the affair. Such contextual factors have, to the best of my knowledge, not been considered in the literature about causal effects of schooling on political attitudes.

The remainder of the paper is structured as follows. Section 1.2 provides details about the study design and balance checks. Section 1.3 presents the main results, section 1.4 investigates potential mechanisms, and section 1.5 describes the so-called *Ibiza affair* and its different effects on apprentices at school and at work. Section 1.6 concludes.

1.2 Study Design

After nine years of compulsory education, students in Austria can decide between continuing with the general educational pathway and starting a vocational education and

training (VET). More than 200 VET professions are offered and a considerable share of the Austrian population receives vocational training. 34% of the Austrian population aged from 25 to 64 held an apprenticeship diploma in 2017 and about the same share of the population starts an apprenticeship each year.[3]

To become an apprentice, one has to apply at a company that offers vocational training. If accepted, he or she receives both training at the company and complementary education at a part-time vocational school. In this dual system, the apprentice is in a training and work relationship with the training company and a student at a vocational school. The company is responsible for providing practical training, which makes up about 80% of the entire training time. In vocational school, apprentices receive general education (languages, civic education and religious education) as well as theoretical and practical education that complements their training. Depending on the profession, the duration of a VET is between two and four years.

I exploit that school classes of the same grade level and profession are as good as randomly assigned to different course blocks, respectively, school terms. In each school year, four course blocks of ten weeks each are offered and it is the vocational school's task to assign its apprentices to one of these terms. Consequently, at the end of the first term one group of apprentices had been at school for ten weeks, whereas apprentices of the second term had been at work in their training companies during that time. Between two school terms we have, therefore, the opportunity to compare groups of apprentices that had been at school with groups of apprentices of the very same grade level and profession that had not been at school during that time period.

I will show that it is reasonable to believe that school classes were as good as randomly assigned to term one (i.e. to receive the "treatment") or to term two (i.e. the "control" condition). Based on that evidence, I infer that at the end of term one and the beginning of term two significant differences between both groups of apprentices in terms of political attitudes can be attributed to the schooling experience.

1.2.1 Assignment of Apprentices to School Terms

The profession of an apprentice and the location of his or her training company determines the vocational school that has to be attended. Within vocational schools, the principals are responsible for assigning their students to one of the four school terms. That has to be done by the end of June for the following school year, which starts at the beginning of September.

In principle, vocational schools are free to assign an apprentice to any term they like.

[3]Statistik Austria (2019) and WKO (2020)

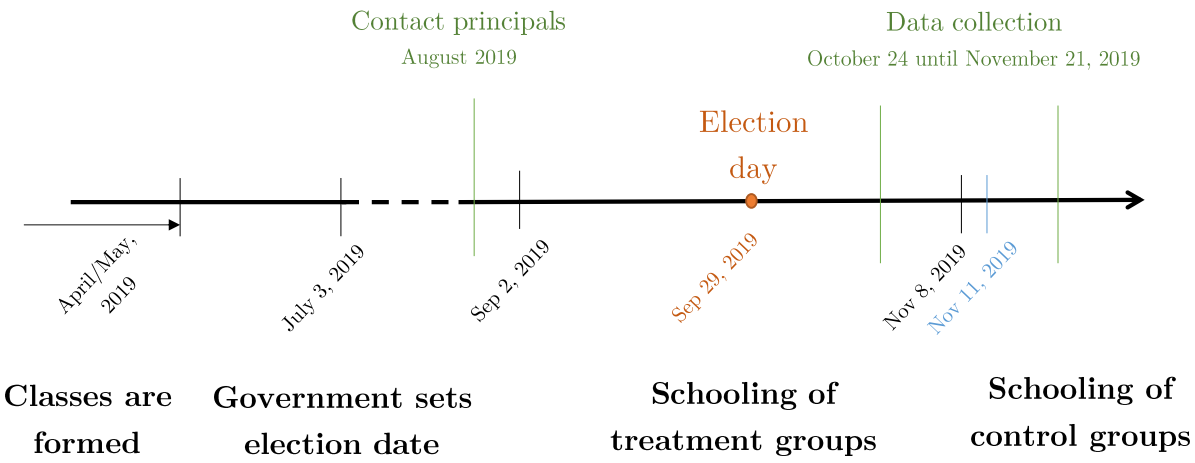


Figure 1.1: Timeline with the most important dates.

Although schools that I contacted claimed that assignment of apprentices to terms was "completely random", they also informed me about some informal guidelines for class assignment:

Some vocational schools base their decisions on the place of residence of the apprentices. This is common for schools that are located in smaller towns that do not have good public transport connections. Apprentices from certain geographic regions are assigned to the same term in order to facilitate carpooling. I ignored schools that base their assignment decisions on the place of residence of apprentices and I did not collect data from any of their classes.

Second, in industries with seasonal fluctuations in workload, schools assign apprentices of the same company evenly across school terms. Apprentices are a valuable part of a company's workforce and schools try to prevent a situation in which some companies lack a large fraction of their apprentices during peak season.

Third, vocational schools usually collaborate very closely with the firms and company requests are fulfilled whenever possible. Some firms allow apprentices to choose their preferred term, others try to distribute their apprentices evenly across terms. But there might be many other reasons for firms to prefer having an apprentice at school for instance in term one and not in term two. It was, of course, not possible to verify the exact reasons for why a single apprentice was sent to term one or to term two by his or her training company.

Several factors make me confident that potential self-selection is not an issue in my setting. I surveyed apprentices of as many as 12 different professions, representing a very diverse set of industries and service sectors. Professions included information technology, office administration, automation technology, plumbing, etc. (see section 1.A in the appendix for the complete list of professions). The data suggests that no particular term

was favored by firms (29.0% of all apprentices in term 1 were sent by their companies compared to 29.9% in term 2) or by individual apprentices (12.8% of apprentices in term 1 and 13.9% in term 2).

For these reasons, I do not expect systematic assignment of apprentices with specific individual characteristics to a specific school term. In section 1.2.4 I provide balance checks that confirm these expectations. Balance tables with only apprentices whose firms suggested a specific school term provide further evidence for non-systematic assignment, and my main findings do not change if I exclude them from the analysis (see section 1.B.1 in the appendix). These indications, therefore, suggest that assuming as good as random assignment of apprentices to terms is justified.

1.2.2 Institutional Background of Vocational Education in Austria

All VET programs are standardized on the national level. Companies that want to provide vocational training are audited by the Chamber of Commerce and by the Chamber of Labor. They check whether the applying firm fulfills all legal (the company has to be entitled to operate in the professional field in which it wants to train apprentices) and operational (the capability of teaching apprentices all the skills that the profession requires) requirements. Furthermore, every company has to have a certified trainer for its apprentices.

The curricula in vocational schools are set by the national government and the federal administrative school boards. The government defines a framework curriculum for each VET program. These framework curricula contain the subjects and their contents that have to be taught as well as the number of total lessons in each of these subjects. The federal administrative school boards then assign the course contents and the number of lessons of each subject to specific grade levels. The VET program content is, therefore, the same in every federal state for a given profession. In which school grade that content is taught can, however, differ between federal states.

The crucial feature of VET programs for my study is that teaching provided by vocational schools can be delivered on a course-specific basis. Apprentices that I focus on are at school during blocks of ten consecutive weeks in each year of their apprenticeship. Other apprentices attend school on a seasonal basis or either once or twice a week during an entire school year, depending on the profession and on the needs of the industry.

1.2.3 Data Collection

Preparatory work started in June 2019, when I called vocational schools in Austria to gather information about their class assignment procedures. Schools had to fulfill certain criteria to be suitable for my study. They had to offer course-specific teaching for at least two classes of apprentices of the same grade level and profession in terms one and two. I checked whether an apprentice's place of residence was considered for class assignment and I disregarded schools that did not confirm its irrelevance. For logistical reasons I also took the locations of vocational schools into consideration. I preferred schools that were close to Vienna and centrally located in their federal states. Schools in Upper Austria and Styria fulfilled above criteria best, so I focused on those two federal states for data collection.

The administrative school boards gave their approval for the research project in mid-August, so I could start contacting the schools at the end of August. Vocational schools in Upper Austria and Styria that were suitable according to the criteria above were asked to participate in the study. Seven out of twelve principals that I contacted showed interest and they allowed me to conduct the surveys in their schools.

In Upper Austria in school year 2019/2020, the first term began on September 2 and ended on November 8, 2019. Apprentices of the second term entered school immediately afterwards, namely on November 11, 2019. In Styria, the school year started a few days later. The first term began on September 9 and ended on November 13, 2019. The second term started on the next day on November 14, 2019.

Between October 24 and November 21, a team of at least two researchers visited the schools and carried out the data collection. Each school was visited twice, once at the end of the first term and once at the beginning of the second term. I surveyed 17 class pairs with one class of the same grade level and profession in both school terms. In total, 34 classes were surveyed – 17 classes in term one and 17 classes in term two. Figure 1.1 shows a timeline with an overview of the most important dates of my study and Table 1.1 provides information on basic background characteristics of apprentices, respectively classes that were surveyed.

The survey contained questions about socio-demographic characteristics, various background variables and three main questions regarding political interest, voting intention and party support. Several of the questions are based on the *Austrian National Election Survey, AUTNES* (Aichholzer et al. 2018), which is a well-established and institutionalized national election study that is designed by political scientists at the University of Vienna. Other questions are based on an unpublished draft questionnaire from *Europe Direct and VieCER (n.d.)*. This questionnaire was designed for an election study at schools, the questions were, therefore, tailored to the needs of young students.

Table 1.1: Summary statistics

	Upper Austria	Styria	Total
Number of schools	6	1	7
Number of classes	28	6	34
Classes with a			
technical profession	18	4	22
commercial profession	10	2	12
Classes in			
grade level two	16	4	20
grade level three	12	2	14
Number of apprentices	657	106	763

Notes. Overview of study participants and class characteristics.

The first two main questions were not incentivized. Apprentices were asked to state their political interest on a scale from 0 (not interested at all) to 10 (very interested) and whether they would turn out to vote in a national election on a scale from 1 (certainly not) to 10 (certainly). Note that the vocational students in my study were 16 years or older, which means that they are entitled to vote in national elections if they are Austrian citizens.[4]

To elicit party support, I offered each apprentice a donation of €10,- to his or her most preferred political party. All parties that participated nationwide in the parliamentary election in 2019 were listed in the survey. *No donation to any party* was also available as an option. Donations were randomly realized with a chance of 20% and paid by the researcher. Apprentices were also informed that I would make the transaction receipts publicly available on the author's website.[5]

The surveys were conducted during school time. Each school had its own computer rooms, so it was not necessary to set up an own lab at the school sites. I equipped each PC in the computer rooms with a mobile privacy screen in order to guarantee anonymous answers for study participants. Apprentices received a public link to the online survey and each of them drew a unique private access code with which they were able to access the survey. The investigators were not able to connect any answer with a

[4]Austria is one of very few countries that have a general voting age of 16 in national elections. In the EU, Malta is the only other country that has also lowered its voting age to 16.

[5]By the end of February 2020 all donations were realized, except to ÖVP. In August 2019 the ÖVP declared that it would not accept any donations from individuals or companies until the end of 2019. They, however, surprisingly continued that policy beyond 2019. I doubt that apprentices could anticipate that when they participated in my survey. Since I promised to realize the donations, I decided to transfer the money to *Hilfswerk Österreich* instead, an Austrian charity organization with strong historical and personnel connections to ÖVP.

study participant. Participation in the study was voluntary and apprentices could leave at any time. Only two out of 769 apprentices decided not to participate in the survey. Data of four participants could not be used for further analyses since three apprentices turned out to be younger than 16 and one apprentice did not answer all questions due to time constraints. All apprentices who participated received a chocolate candy as a small thank-you gift.

1.2.4 Balance Checks

To test whether it is plausible to assume that classes were as good as randomly assigned to term one or to term two, I estimate the following linear probability model with class pair fixed effects:

$$x_{ik} = \alpha_k + T_{ik}\boldsymbol{\beta} + e_{ik} \tag{1.1}$$

$k = 1, ..., 17$ denotes the class pairs with $i = 1, ..., n_k$ individuals in each class pair. x_{ik} is the variable for which randomization is checked. α_k is the fixed effect for class pair k, T_{ik} is the treatment indicator for person i in class pair k, and β is the unknown parameter. e_{ik} is an individual specific error term, which is assumed to be independent between class pairs.

We can interpret the coefficient β as the difference between x_{ik} in the control and treatment groups after adjustments for different class sizes. Statistically significant differences would indicate that the assumption of as good as random assignment of classes into treatment and control has to be rejected. The data is clustered on the level of class pairs, which is the level of sampling conditional on the study approval from a principal (Abadie et al. 2022, Chaisemartin and Ramirez-Cuellar 2023). With just 17 class pairs, standard adjustments for clustering can lead to downward biased cluster-robust standard errors (Donald and Lang 2007). I, therefore, report p-values and 95% confidence intervals based on the wild cluster bootstrap-t method (Cameron, Gelbach, and Miller 2008, Roodman et al. 2019).

Variables for the balance checks were pre-registered and selected based on their potential influence on any outcome variable and on their capability to act as a proxy for an influential (but unobservable) characteristic. Furthermore, it was required that the variables used could not be affected by ten weeks of schooling or working.

The balance tables look promising. Table 1.2 shows summary statistics for variables with binary and numerical data. We observe one estimate (*grade point average, GPA*) that is statistically significant at the 10% level, which is what we expect to find under randomization given twelve variables. Table 1.3 provides balance statistics for variables

Table 1.2: Balance table I

	Control Groups		Treatment Groups		Test for Difference		
	Mean	SD	Mean	SD	β	95% CI	p-value
Age	18.11	2.48	18.46	2.58	0.283	[-0.296, 0.736]	0.375
Female	0.39	0.49	0.42	0.49	0.024	[-0.028, 0.077]	0.365
Austrian citizenship	0.89	0.31	0.88	0.33	-0.016	[-0.060, 0.031]	0.466
First apprenticeship	0.95	0.21	0.95	0.23	-0.005	[-0.045, 0.036]	0.813
Work experience†	0.49	1.10	0.54	1.13	0.048	[-0.095, 0.191]	0.483
Income†	5.49	1.38	5.37	1.42	-0.123	[-0.337, 0.081]	0.234
Exempted from							
civic education	0.22	0.41	0.19	0.39	-0.033	[-0.084, 0.020]	0.197
Attend boarding school	0.30	0.46	0.36	0.48	0.055	[-0.061, 0.172]	0.329
Grade point average	1.95	0.78	2.13	0.85	0.177*	[-0.008, 0.358]	0.060
Trade union member†	0.29	0.46	0.34	0.47	0.027	[-0.072, 0.126]	0.577
Active in a political party†	0.07	0.26	0.09	0.28	0.014	[-0.030, 0.058]	0.497
Vocational training							
started in summer	0.75	0.43	0.74	0.44	-0.019	[-0.108, 0.069]	0.649
N	411		352				

†Less observations due to non-responses (Income: $N = 741$, Active in a union: $N = 617$, Active in a party: $N = 658$). Income was measured in steps of €150. *5* denotes net income of €750 to €900 and *6* denotes €900 to €1050, for example.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

with categorical data. None of the joint $F-$tests for these variables shows significance at a conventional significance level. However, students from training companies with 10-49 employees, and students who prepare for the *A-levels* ("Matura") are not evenly distributed in my sample. Two slightly imbalanced categories are again not surprising given the 18 different categories in my analysis. I, nevertheless, include these variables in my regression analyses as additional covariate controls. We will see that this increases the treatment effect estimates, which means that those imbalances render the estimates in my baseline analyses more conservative.

Table 1.3: Balance Table II

	Control Groups		Treatment Groups		Test for Difference		
	N	%	*N*	%	β	95% CI	p-value
Company size							**0.274**
0-9 employees	60	14.60	47	13.35	-0.001	[-0.040, 0.036]	0.964
10-49 employees	109	26.52	118	33.52	0.062*	[-0.003, 0.126]	0.059
50-249 employees	66	16.06	57	16.19	-0.000	[-0.066, 0.061]	0.989
250 employees or more	176	42.82	130	36.93	-0.061	[-0.143, 0.021]	0.129
Hometown							**0.709**
<5.000 inhabitants	228	55.47	201	57.10	0.021	[-0.076, 0.121]	0.646
5.000 to 20.000	81	19.71	73	20.74	0.008	[-0.048, 0.061]	0.762
20.000 to 100.000	44	10.71	40	11.36	0.006	[-0.038, 0.050]	0.766
>100.000 inhabitants:							
Center of a city	33	8.03	22	6.25	-0.019	[-0.061, 0.023]	0.349
Outskirts of a city	25	6.08	16	4.55	-0.017	[-0.044, 0.012]	0.220
A-levels (Matura)							**0.248**
Have Matura degree	21	5.11	24	6.82	0.011	[-0.064, 0.073]	0.820
Prepare for Matura	89	21.65	55	15.63	-0.066*	[-0.138, 0.006]	0.069
Neither have Matura nor prepare for it	301	73.24	273	77.56	0.055	[-0.042, 0.157]	0.267
N	411		352				

Notes. For the joint significance tests (p-values in bolt), the dummy for *treatment assignment* was regressed on the categorical variable.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.3 Results

In this section I report estimates for the effects of schooling on political interest, voting intention and party support. I use the following simple linear probability model to analyze my data.[6]

$$y_{ijk} = \alpha_k + \mathbf{x'}_{ijk}\boldsymbol{\gamma} + T_{jk}\delta + r_{jk} + e_{ijk}. \tag{1.2}$$

y_{ijk} represents the response of person i in class j of class pair k. The class pair specific intercepts α_k are modeled as fixed effects, $\mathbf{x'}_{ijk}$ is a vector of covariates, T_{jk} is a treatment indicator and δ denotes the mean treatment effect, which is my main variable of interest. r_{jk} and e_{ijk} are class-specific and individual-specific error terms.

Data are clustered on the level of class pairs, and p-values and confidence interval were calculated based on the wild cluster bootstrap-t method. I estimate four models with different sets of controls for each of the three outcome variables: Political interest, voting intention and party support. Model (1) does not include any covariates. To increase precision of estimates, my baseline model (2) controls for *age* and *gender* of apprentices. Models (3) and (4) include additional covariates that appeared to be imbalanced in the randomization checks: Model (3) adds *GPA* and model (4) includes categorical covariate controls for *size of training company* and *A-levels* (which indicates whether the apprentice has the Matura degree, prepares for it or none of these).

In the following sections I report the effects of schooling on political interest, voting intentions, and party support of vocational students.

1.3.1 Political Interest

Table 1.4 shows my estimates of the effect of schooling on political interest. The treatment indicator is positive and statistically significant in any of my model specifications. The point estimate in my main specification suggests that the period of ten weeks in school increased political interest of apprentices by 0.586 on the eleven-point scale, which is an increase of 0.22 standard deviations.

Controlling for variables that appeared to be slightly imbalanced between treatment and control groups increases the treatment effect estimates. The estimate in my main specification is, therefore, conservative due to these imbalances.

[6] I do not use a SURE model for two reasons. First, point estimates would not differ because each set of equations contains the same set of regressors. Second, I have to use the wild cluster bootstrap-t method to calculate p-values and 95% confidence intervals. Because of the low number of clusters in my data, standard adjustments for clustering would lead to downward biased cluster-robust standard errors, as also noted in section 1.2.4.

Table 1.4: Effects of schooling on political interest

	(1)	(2)	(3)	(4)
Treated	0.565**	0.586**	0.710**	0.740**
	[0.005, 1.119]	[0.027, 1.130]	[0.158, 1.250]	[0.197, 1.271]
Age		-0.020	-0.050**	-0.071***
		[-0.101, 0.043]	[-0.139, -0.005]	[-0.204, -0.006]
Female		-0.664**	-0.783***	-0.776***
		[-1.195, -0.140]	[-1.281, -0.295]	[-1.294, -0.262]
Grade point average	No	No	Yes	Yes
Additional controls	No	No	No	Yes
N (apprentices)	763	763	763	763
N (class pairs)	17	17	17	17

Notes. Dependent variable is *political interest*, from 0 (not interested at all) to 10 (very interested). 95% confidence intervals in brackets. Additional controls account for *size of training company* and *A-levels*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As commonly observed in surveys on political interest, women in my sample also stated lower levels of political interest than men (Bennett and Bennett 1989, Verba, Burns, and Schlozman 1997).

1.3.2 Voting Intention

Education is consistently found as one the strongest predictors of voter turnout (Smets and Ham 2013). A common explanation for this relationship is the increase in political interest through education. Dostie-Goulet (2009) and Miklikowska, Rekker, and Kudrnac (2022) show that political discussions in school increase students' political interest. And people who are more interested in politics are, in turn, more likely to be politically active (Milbrath and Goel (1977, p. 46) even write that "[...]the relationship is so regular that many authors do not bother to report it[...]").

However, despite the significant increase in political interest that was directly caused by schooling, my results on the effect of schooling on voting intention are inconclusive. Table 1.5 shows the regression table for the effect of schooling on *voting intention* as the dependent variable.[7] We see that treatment effects are all positive but not significant in any of my model specifications.

[7] *Voting intention* was elicited on a scale from 1 (no intention) to 10 (strong intention) in order to exclude indecisive answers (i.e. the mean value) by design.

Table 1.5: Effects of schooling on voting intention

	(1)	(2)	(3)	(4)
Treated	0.142	0.167	0.331	0.339
	[-0.264, 0.559]	[-0.228, 0.573]	[-0.068, 0.748]	[-0.110, 0.806]
Age		-0.057	-0.096*	-0.101*
		[-0.184, 0.064]	[-0.220, 0.006]	[-0.258, 0.017]
Female		-0.359	-0.517*	-0.486
		[-1.020, 0.240]	[-1.164, 0.052]	[-1.171, 0.108]
Grade point average	No	No	Yes	Yes
Additional controls	No	No	No	Yes
N (apprentices)	763	763	763	763
N (class pairs)	17	17	17	17

Notes. Dependent variable is *voting intention*, from 1 (no intention) to 10 (strong intention). 95% confidence intervals in brackets. Additional controls account for *size of training company* and *A-levels*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.3.3 Party Support

Table 1.6 reports estimates of the direct effects of schooling on the support for a specific political party. Every entry in the table shows the estimated effect of schooling on the likelihood to make a financial donation to the party indicated in the first column.

I do not find evidence for an effect of additional schooling on party support of vocational students. In my baseline specification, I only observe a marginally significant positive effect of schooling on the likelihood to support the conservative *Austrian People's Party* (ÖVP). After including *grade point average* as an additional control variable, the treatment effect estimate increases and becomes significant at the 5% level. None of the other estimates is significant at conventional significance levels.

Table 1.6: Effects of schooling on the likelihood to support a certain political party

Party	Sample Mean	(1)	(2)	(3)	(4)
ÖVP	0.165	0.041	0.048*	0.057**	0.055**
		[-0.019, 0.101]	[-0.009, 0.104]	[0.004, 0.111]	[0.003, 0.108]
SPÖ	0.152	-0.013	-0.011	-0.020	-0.017
		[-0.071, 0.042]	[-0.068, 0.043]	[-0.074, 0.032]	[-0.069, 0.068]
FPÖ	0.189	-0.027	-0.024	-0.031	-0.036
		[-0.113, 0.064]	[-0.104, 0.065]	[-0.116, 0.061]	[-0.119, 0.056]
GREENS	0.157	0.013	0.009	0.012	0.011
		[-0.033, 0.059]	[-0.031, 0.050]	[-0.028, 0.052]	[-0.027, 0.052]
NEOS	0.058	0.019	0.018	0.024	0.026
		[-0.015, 0.051]	[-0.013, 0.059]	[-0.009, 0.054]	[-0.007, 0.054]
Other Parties	0.025	0.013	0.001	0.001	0.000
		[-0.023, 0.026]	[-0.023, 0.025]	[-0.024, 0.026]	[-0.025, 0.023]
No Party	0.254	-0.033	-0.040	-0.043	-0.039
		[-0.098, 0.032]	[-0.104, 0.024]	[-0.108, 0.022]	[-0.108, 0.026]
Included Covariates					
Age and gender		No	Yes	Yes	Yes
Grade point average		No	No	Yes	Yes
Additional controls		No	No	No	Yes
N (apprentices)		763	763	763	763
N (class pairs)		17	17	17	17

Notes. Each entry shows the estimate of the effect of schooling on the likelihood to support a certain political party. 95% confidence intervals in brackets. Additional controls account for *size of training company* and *A-levels*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.4 Potential Mechanisms

With my rich survey data I can test several potential mechanisms for the effects of schooling that are proposed in the literature. First, I test the social network hypothesis and the cognitive ability hypothesis. Second, I investigate whether knowledge about the positions of Austrian parties increases and whether the quality of vote improves due to schooling. The statistical analyses in this section are all based on model 1.2 with the baseline covariates *age* and *gender* if not indicated otherwise.

1.4.1 Social Network Hypothesis and Cognitive Ability Hypothesis

The two most prominent hypotheses for why schooling could affect political attitudes of citizens are the social network hypothesis and the cognitive ability (or civic education) hypothesis. My research design allows to investigate both hypotheses.

According to the social network hypothesis it is the peers in school who affect political attitudes of apprentices. They might discuss political issues which could raise political interest, provide each other with political information or exert social pressure to motivate others to turn out to vote (Funk 2010, Dellavigna et al. 2017). The cognitive hypothesis states that schooling enhances the cognitive skills that citizens need to effectively participate in democracies (Wolfinger and Rosenstone 1980). Students learn how to participate in the political process and how to gather information about the political parties in their country. Such civic competencies increase political participation and they improve the quality of citizens' decisions (see Solis (2017) for a discussion).

To test the social network hypothesis, I exploit that some of the apprentices spend much more time with their peers than others. For each profession, only a very small number of specialized vocational schools exists. They are usually located in the major towns of each federal state, which means that a sizable share of apprentices cannot commute to school each day due to long travel distances. As an alternative, they stay in a boarding school with their peers (33% of all apprentices in my sample). To estimate the effects of staying in a boarding school and, hence to have more contact with peers from school I use baseline model 1.2 with two additional dummy variables: One that indicates whether an apprentice attends boarding school or not and the other is its interaction with the treatment indicator. All p-values below refer to the estimated effect of the latter interaction term. I do not observe any significant effect of attending boarding school on apprentices' political interest ($p = 0.665$), voting intention ($p = 0.774$), or support for a specific political party (effect on support for *ÖVP*, $p = 0.617$; *SPÖ*, $p = 0.850$; *FPÖ*, $p = 0.942$; *GREENS*, $p = 0.658$; *NEOS*, $p = 0.771$; *no party*, $p = 0.727$).

For the cognitive ability hypothesis I can find a weak indication in the data. Vocational schools in Austria offer *civic education*, which is a compulsory subject[8] that is specifically designed to improve the civic competencies of apprentices. I exploit that the number of *civic education* lessons in the curriculum varies across vocations, school grades, and federal states. To estimate the effects of *civic education* I proceed analogous to the analysis above. I use baseline model 1.2 with two additional variables, one for the number of civic education lessons that an apprentice receives per week and its interaction with the treatment indicator. Again, p-values below refer to the estimated effect of the latter interaction term. There is no significant effect of the number of lessons in *civic education* on apprentices' political interest ($p = 0.461$) or on voting intention ($p = 0.371$). Only support for the ÖVP seems to increase by 0.004 on average for each lesson of *civic education* per week, $p = 0.0409$ (effect on support for SPÖ, $p = 0.5681$; FPÖ, $p = 0.5139$; GREENS, $p = 0.5951$; NEOS, $p = 0.5897$; *no party*, $p = 0.9743$). However, due to the multiple hypothesis testing problem, the interpretation of one significant result remains ambiguous.

1.4.2 Knowledge about Politics

The social network hypothesis and the cognitive skill hypothesis both suggest that schooling improves citizens' knowledge about politics. With my data I can test whether schooling affects knowledge about the positions of political parties. In Austria an editorial team of journalists and political scientists summarizes the political positions of Austrian parties before national elections in the following way. They work out a list of questions that cover a broad range of political topics. Then they ask each party that stands for election to answer each of the questions with either "yes" or "no". In addition, the parties are asked to weight each question according to its importance for the party on a scale from one (less relevant) to three (very important). If an answer or a weighting does not match the policy of a party, the editorial team can make adaptions. About six weeks before the election the list of questions is published in form of an election compass on the website `https://wahlkabine.at/` (subsequently denoted as *Wahlkabine*).[9] Interested citizens can then answer the questions to see how their political options match the positions of the political parties that stand for election.

For five of these questions I asked the apprentices to guess the answers that the five parties with the highest shares of votes had given. To make certain terms easier

[8]If an apprentice has already had *civic education* in school, he or she can be exempted from this subject by the principal upon request. About 20% of apprentices in my sample were exempted from *civic education*.

[9]See `https://wahlkabine.at/nationalratswahl-2019/stellungnahmen` for the list of all questions and the positions of the political parties in 2019. Retrieved on January 27, 2022.

understandable, the additional information that *Wahlkabine* provides online was also shown in my survey (see the footnotes below). The following questions were shown in my survey.

And what do the parties think about the following question:

Q38 Should the weekly working time be reduced to 30 hours without wage adjustment?
Q39 Should company taxes[10] *be reduced?*
Q40 Should the joint school for 10 to 14 year-olds be realized?
Q41 Should criminal foreigners be deported, independent of their duration of stay in Austria?
Q42 Should a CO2-tax[11] *be introduced?*

Each question had to be answered for each of the five political parties with the highest shares of votes in Austria: ÖVP, SPÖ, FPÖ, GREENS, and NEOS. The apprentices had to select one of the five answers below.

1. I am sure: the party agrees
2. I believe: this party agrees
3. I don't know
4. I believe: the party disagrees
5. I am sure: the party disagrees

If schooling increased knowledge about party positions, then apprentices who went to school should have been able to answer more questions correctly than those apprentices who were at work. However, I do not find evidence for such an effect. The average number of correct answers is slightly higher in the group of apprentices that went to school (+0.38) but the difference is not significant ($p = 0.368$). In a second specification I weight the answers by the apprentice's confidence: An answer is counted as +2 (-2) if the apprentice was right (wrong) and sure about his or her answer, +1 (-1) if the apprentice was right (wrong) and not sure about his or her answer, and 0 if the apprentice did not know. The effect of schooling on this weighted outcome variable is +0.33 but again insignificant ($p = 0.593$).

[10] Company taxes refer to all taxes that companies have to pay. These taxes include profit taxes (like income tax, corporate income tax, and business tax), consumption taxation (value added tax and real estate transfer tax), and property taxation (land tax, inheritance tax and gift tax).
[11] A CO2-tax taxes carbon dioxide and other greenhouse gases that substantially contribute to the climate change crisis. The subject of taxation is mainly the combustion of fossil energy sources, for example in the industry, but also fuels.

1.4.3 The Match between Political Opinion and Party Support

To assess whether schooling improves the match between an apprentice's political opinions and the positions of the political party that he or she supports, I asked apprentices to answer 13 questions from *Wahlkabine*. The questions cover a broad range of political topics, see Appendix 1.A.2 for the list of questions that I used. Apprentices could answer with "yes", "no", or "no answer". In addition, the apprentices were asked to rate the importance of each question from 1 (less relevant) to 9 (very important).

I analyze whether schooling increases the number of answers that coincide with the political positions of the party that an apprentice supports. For the analysis I also construct a weighted measure of how well the political opinions of an apprentice matches the positions of his or her supported party. This measure is identical to the measure that is used by *Wahlkabine* to assess the closeness of one's political opinions and the positions of the political parties that stand for election.[12] The measure is calculated as the product sum of the apprentice's and the party's weight on the respective question and the factor +1 if the answer of the apprentice and the party coincide, -1 if they differ and 0 if the apprentice did not answer the question.

I do not find evidence for an effect of schooling on the quality of match between an apprentice's political opinions and his or her support for a party. Neither the unweighted number of coinciding answers increases significantly (+0.11, $p = 0.627$) nor the weighted number of coinciding answers (-2.10, $p = 0.584$).

1.4.4 Teachers and Peers

Political discussions with parents, friends or teachers play a crucial role in fostering a student's interest in politics. Dostie-Goulet (2009) shows that an increase in the frequency of political discussions in teenagers' social networks leads to an increase in their political interest. Miklikowska, Rekker, and Kudrnac (2022) confirms this positive relationship between political discussions in school and political interest of students. When students have the impression that teachers try to involve them in discussions about political issues, their general interest in politics increases.

Engagement in political discussions with teachers and with fellow students has, therefore, very likely also affected apprentices in my study. In school, apprentices meet teachers who might be more enthusiastic about politics than their regular workmates. They might engage in political discussions with them or with fellow students, which could stimulate their own political interest. We indeed find suggestive evidence for this mechanism in our

[12] See `https://wahlkabine.at/downloads/Wahlkabine_Methodik_NRW2019.pdf`, retrieved on January 27, 2022.

data. It seems that teachers were able to pass on their political interest to students.[13]

In my study, I asked apprentices how often they had talked about politics with different people or groups. Not surprisingly, apprentices in school talked more often with teachers ($p = 0.000$) and fellow students ($p = 0.000$) about politics and less often with workmates ($p = 0.006$). The frequency of discussions with others did not change significantly.[14]

Table 1.2 shows how politically interested these reference groups appear to apprentices on a scale from 0 (not at all interested) to 10 (very interested). We see that teachers are perceived as the most interested group (7.05 on average), classmates are perceived as being least interested in politics (4.49 on average) and the perceived political interest of workmates is in between (5.80 on average).[15]

It seems, therefore, plausible to conjecture that political discussions with politically highly interested teachers were a driving factor for the increase in political interest of apprentices.

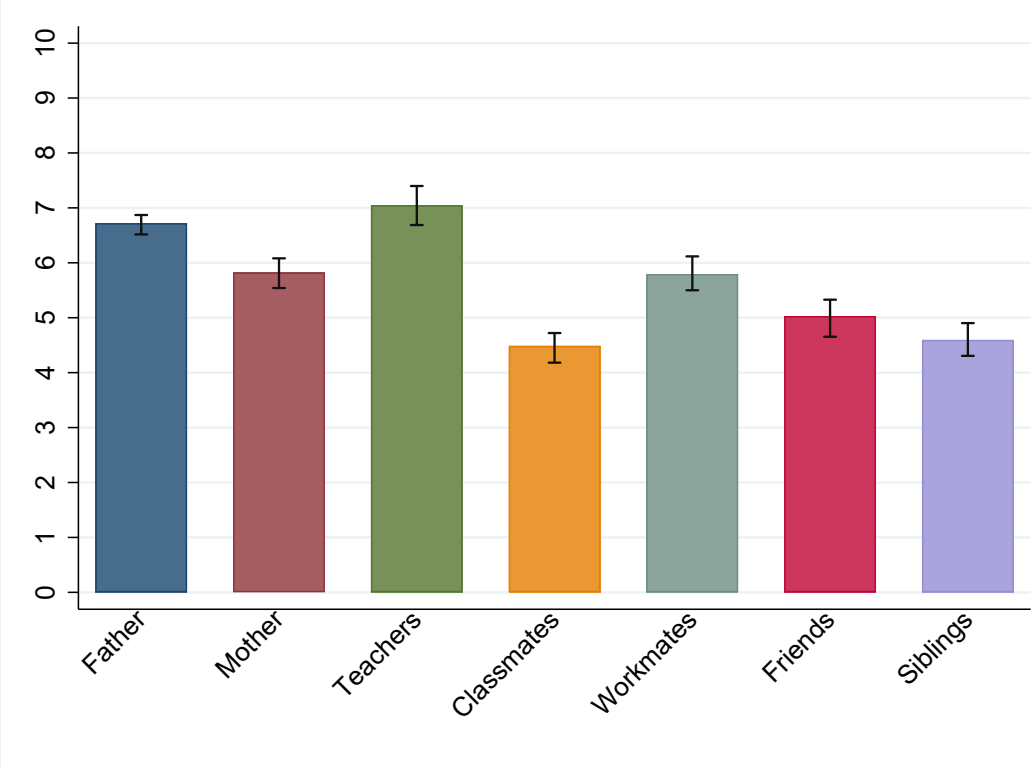


Figure 1.2: How politically interested others appear on a scale from 0 (not at all interested) to 10 (very interested). Error bars represent 95% confidence intervals.

[13]The analyses in this subsection were not pre-registered.
[14]Father ($p = 0.652$), mother ($p = 0.132$), siblings ($p = 0.771$), friends ($p = 0.541$).
[15]All three values are significantly different from each other ($p = 0.000$ for each comparison, unpaired t-tests between class pairs' mean values of each variable).

1.5 The General Political Situation Matters

My research design allows to investigate whether the general political situation in which schooling happens matters.[16] After every school year, vocational students have been in school for the very same number of weeks. The specific school term that they were assigned to, however, differs. I find that vocational students who attended school during exposure of a major political scandal in Austria were significantly less likely to support the party whose officials were responsible for the affair.

On May 17, 2019, a secretly recorded video was published, in which two high-level officials of the *Austrian Freedom Party* (FPÖ) were, inter alia, allegedly promising public contracts in return for campaign support. The FPÖ, an Austrian right-wing populist party with immigration as its core issue (Ennser-Jedenastik 2016), received almost 26% of the votes in the national election 2017. Since then it had been the junior partner in a coalition with the conservative *Austrian People's Party* (ÖVP). The so-called *Ibiza affair* (named after the island where the video was shot) eventually caused the collapse of this governing coalition and it led to snap elections in September, 2019.

The fourth term in that school year 2018/19 lasted from April 8 until June 28, 2019. Hence, when the video was published and publicly discussed, 26% of the apprentices in my sample happened to be at school whereas the others were working in their training companies.

To test whether apprentices in school were differently affected by the *Ibiza affair* than those who were working, I apply equation 1.2 and use schooling in term 4 of the previous school year as the treatment indicator. Table 1.7 summarizes my estimation results. In each row I report the point estimates of the effects of schooling in term 4 of the previous school year on the likelihood to support the party in the first column. I ran five regressions for each party with different sets of control variables. *Additional schooling* is *one* if an apprentice had been at school in the ten weeks before I conducted the survey and *zero* otherwise. This variable is added as a control variable in models two to five.

We see that apprentices who were at school during exposure of the *Ibiza affair* were five months later (when I conducted the survey) significantly less likely to support FPÖ and more likely to support no party at all. There is no significant difference for any other party, neither for schooling in term 4 nor for any other term of the previous school year. If I include a term for the interaction between *additional schooling* and *schooling during exposure of the Ibiza affair*, the negative effect on support for the FPÖ in my main specification is still significant. However, randomization checks (see Appendix 1.C) show slight imbalances between treated and untreated apprentices. My results do not change

[16] The analyses in this section were not pre-registered.

Table 1.7: Effects of schooling during the *Ibiza affair* on the likelihood to support a specific political party

	(1)	(2)	(3)	(4)	(5)
ÖVP	0.018	0.032	0.039	0.033	0.027
	[-0.051, 0.085]	[-0.040, 0.105]	[-0.034, 0.113]	[-0.040, 0.105]	[-0.044, 0.094]
SPÖ	-0.044*	-0.051	-0.050	-0.045	-0.033
	[-0.102, 0.008]	[-0.118, 0.011]	[-0.118, 0.010]	[-0.110, 0.013]	[-0.091, 0.018]
FPÖ	-0.059**	-0.071**	-0.065**	-0.061**	-0.069**
	[-0.117, -0.008]	[-0.132, -0.013]	[-0.121, -0.012]	[-0.113, -0.011]	[-0.119, -0.019]
GREENS	0.016	0.021	0.021	0.019	0.015
	[-0.054, 0.088]	[-0.058, 0.098]	[-0.055, 0.097]	[-0.057, 0.096]	[-0.060, 0.092]
NEOS	-0.006	-0.000	-0.003	-0.006	-0.003
	[-0.063, 0.056]	[-0.055, 0.063]	[-0.060, 0.062]	[-0.065, 0.061]	[-0.059, 0.064]
Other Parties	-0.008	-0.008	-0.010	-0.010	-0.011
	[-0.023, 0.010]	[-0.026, 0.010]	[-0.030, 0.010]	[-0.031, 0.011]	[-0.037, 0.014]
No Party	0.083**	0.077**	0.068**	0.071**	0.074**
	[0.011, 0.160]	[0.009, 0.151]	[0.001, 0.139]	[0.004, 0.140]	[0.006, 0.143]
Included Covariates					
Additional schooling	No	Yes	Yes	Yes	Yes
Age and gender	No	No	Yes	Yes	Yes
Grade point average	No	No	No	Yes	Yes
Additional controls	No	No	No	No	Yes
N (apprentices)	763	763	763	763	763
N (class pairs)	17	17	17	17	17

Notes. Each entry shows the estimate of the effect of schooling in term 4 in 2018/19 on the likelihood to support a certain political party. 95% confidence intervals in brackets. Additional controls account for the apprentice's *citizenship*, the *size of hometown*, whether it is the *first apprenticeship* and whether the *vocational training started in summer*.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

qualitatively when I control for these imbalances.

1.6 Conclusion

A large scientific literature documents the effects of schooling on political preferences, voter turnout, and political interest. In this paper I use a quasi-experimental design to study whether it is schooling *per se* that affects political attitudes of apprentices. For identification I exploit that vocational students in Austria are as good as randomly assigned to different classes and school terms, respectively. This allows comparing the political attitudes of vocational students who only differ in their period of schooling but not with respect to other characteristics. This article, therefore, provides a novel approach to study the direct effects of schooling. One advantage of the research design is that it can be repeated in each school year. Extending the scope of research and addressing other research questions as well seems a natural next step for future research on the effects of schooling.

In my study I find that apprentices who attended school are significantly more inter-

ested in politics. This increase in political interest seems to be driven by political discussions with teachers who were perceived as highly interested in politics. The increase in political interest is, however, not accompanied by a significant increase in voting intention or by a change in support for a specific political party. For a policy maker these findings suggest that political interest of students can be increased by teachers who are themselves highly interested in politics. This factor could, therefore, be considered in teacher training programs or in the hiring process of prospective teachers.

Furthermore, the uncovering of the *Ibiza affair* in Spring 2019 allowed to test whether the general political situation matters for the effects of schooling. And indeed, apprentices who were at school and not at work during exposure of the affair were five months later still less likely to support the FPÖ and more likely to support no party at all. Ten weeks of schooling can, therefore, already affect, which party an apprentice supports. However, it seems that a very strong trigger like the *Ibiza affair* is necessary for this effect to evolve during the relatively short period of schooling that I consider in this study. Studying the interaction effects of schooling and such contextual factors still seems a fruitful opportunity for future research.

1.A Additional Information, Graphs and Tables

1.A.1 List of Professions

I surveyed apprentices from the following professions:

1. computer science
2. plumbing
3. office administration
4. retail business
5. automation technology
6. pharmaceutical assistance
7. electrical engineering
8. mechatronics
9. construction technology
10. metal technology
11. chemical engineering
12. coating technology

1.A.2 Political Opinions

I asked apprentices about their political opinion on the following 13 questions.

Now we focus on political topics. What do you think about the following questions?

Q25 Should the weekly working time be reduced to 30 hours without wage adjustment?
Q26 Should company taxes be reduced?
Q27 Should inheriting remain tax-exempt?
Q28 Should the legal retirement age be raised?
Q29 Should an independent agency be established that investigates police violence?
Q30 Should the joint school for 10 to 14 year-olds be realized?
Q31 Should asylum seekers get a work permit?
Q32 Should criminal foreigners be deported, independent of their duration of stay in Austria?
Q33 Should the general right to vote in national elections remain restricted to Austrian citizens?
Q34 Should the Court of Auditors be allowed to audit all party finances and to impose penalties?
Q35 Should the compulsory broadcasting fees („GIS Gebühren“) be abolished?
Q36 Should internet forums be obliged to record the real names of all its users and to release them if necessary?
Q37 Should a CO2-tax be introduced?

Apprentices could answer with *“yes”*, *“no”*, or *“no answer”*. In addition, the apprentices were asked to rate the importance of each question from 1 (less relevant) to 9 (very

important).

1.A.3 Histograms - Political Interest and Voting Intentions

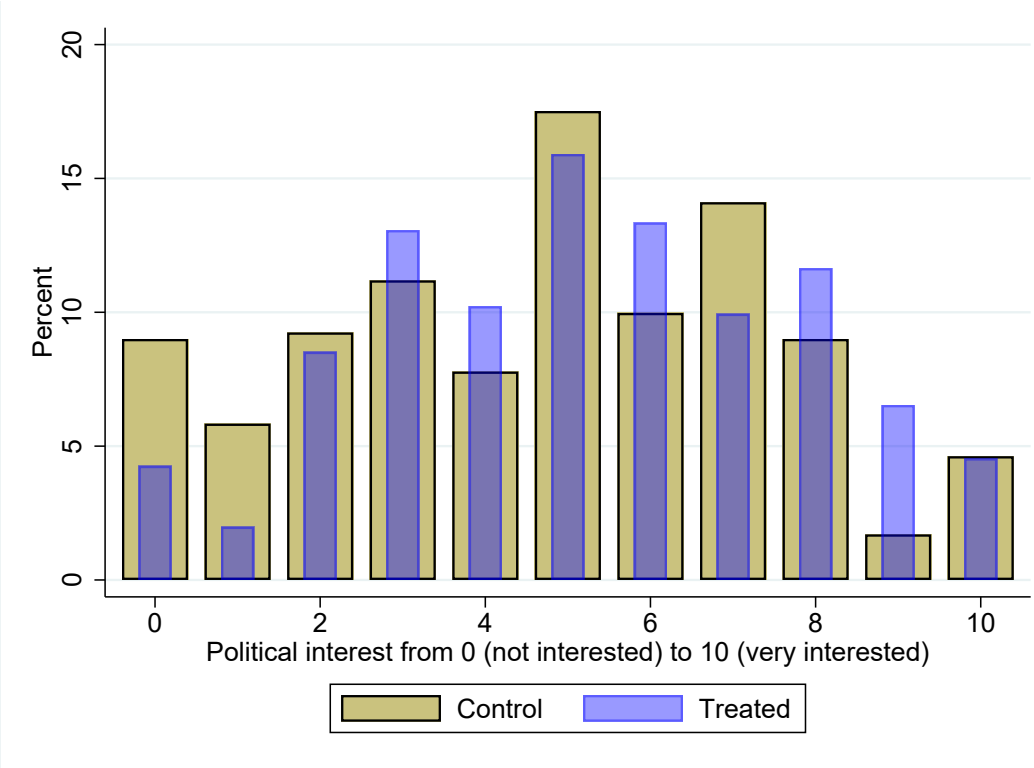


Figure A.3: Histogram of political interest

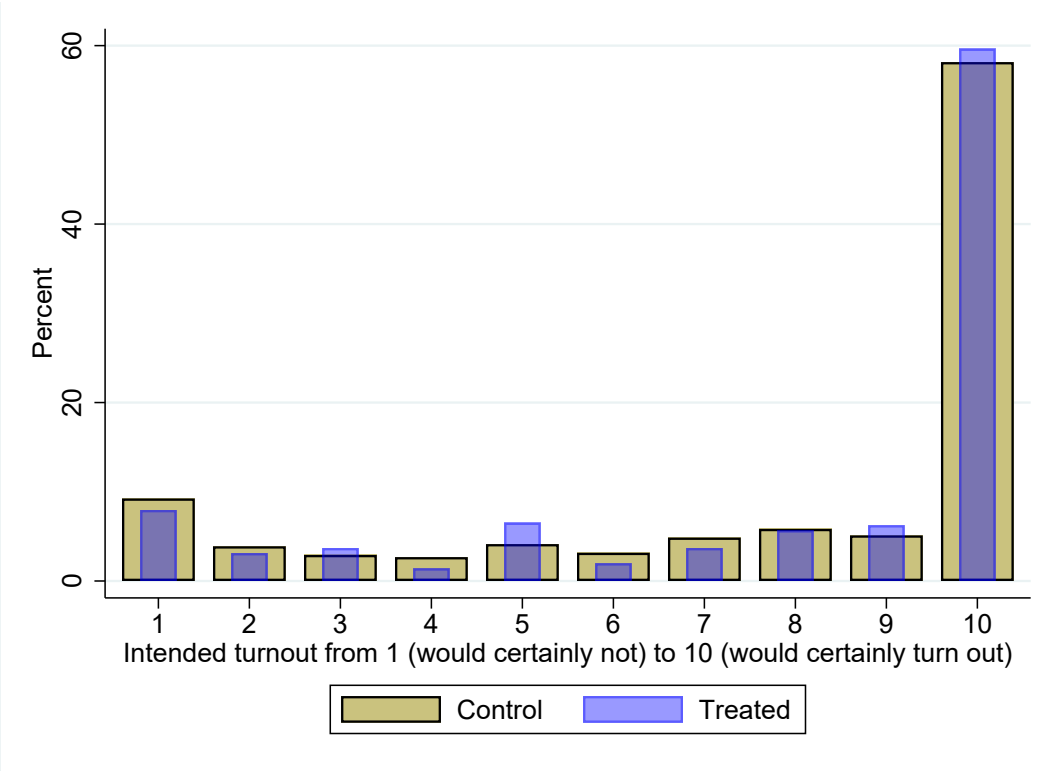


Figure A.4: Histogram of voting intentions

1.A.4 Additional Summary Statistics

Table A.8: Summary statistics for party support

	Control groups		Treatment groups		Total		Parliamentary
	N	%	N	%	N	%	election 2019
ÖVP	60	14.6%	66	18.8%	126	16.5%	28.0%
SPÖ	65	15.8%	51	14.5%	116	15.2%	15.8%
FPÖ	84	20.4%	60	17.0%	144	18.9%	12.1%
GREENS	61	14.8%	59	16.8%	120	15.7%	10.4%
NEOS	20	4.9%	24	6.8%	44	5.8%	6.1%
Other	10	2.4%	9	2.6%	19	2.5%	2.4%
Nonvoters	111	27.0%	83	23.6%	194	25.4%	25.3%

Notes. This table shows the share of apprentices that support a specific party in my sample. The result of the parliamentary election 2019 is shown in the last column.

1.A.5 Standardized Treatment Effects

Table A.9: Standardized treatment effect on political interest

	(1)	(2)	(3)	(4)
Treated	0.213**	0.221**	0.268**	0.279**
	[0.002, 0.422]	[0.010, 0.426]	[0.060, 0.472]	[0.074, 0.479]
Age		-0.008	-0.019**	-0.027***
		[-0.038, 0.016]	[-0.052, -0.002]	[-0.077, -0.002]
Female		-0.250**	-0.295***	-0.293***
		[-0.451, -0.053]	[-0.483, -0.111]	[-0.488, -0.099]
Grade point average	No	No	Yes	Yes
Additional controls	No	No	No	Yes
N (apprentices)	763	763	763	763
N (class pairs)	17	17	17	17

Notes. Dependent variable is *political interest*, standardized with mean 0 and variance 1. 95% confidence intervals in brackets. Additional controls account for *size of training company* and *A-levels*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: Standardized treatment effect on voting intention

	(1)	(2)	(3)	(4)
Treated	0.046	0.053	0.106	0.109
	[-0.085, 0.180]	[-0.073, 0.184]	[-0.022, 0.240]	[-0.035, 0.259]
Age		-0.018	-0.031*	-0.032*
		[-0.059, 0.021]	[-0.071, 0.002]	[-0.083, 0.005]
Female		-0.115	-0.166*	-0.156
		[-0.328, 0.077]	[-0.374, 0.017]	[-0.376, 0.035]
Grade point average	No	No	Yes	Yes
Additional controls	No	No	No	Yes
N (apprentices)	763	763	763	763
N (class pairs)	17	17	17	17

Notes. Dependent variable is *voting intention*, standardized with mean 0 and variance 1. 95% confidence intervals in brackets. Additional controls account for *size of training company* and *A-levels*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.B Robustness Checks

1.B.1 Self-selection

As noted in section 1.2.1, firms can propose a specific school term for their apprentices and schools usually follow these suggestions. Because of this self-selection, my assumption of as good as random assignment of apprentices into treatment and control classes could be violated.

Several observations make me confident that my main results are not caused by such self-selection issues. First, no term was particularly favored by firms or apprentices. As already mentioned above, 29.0% of all apprentices were sent to term 1 because of their companies' preference compared to 29.9% in term 2 ($p = 0.811$, Fisher's exact test). 12.8% of apprentices selected term 1 and 13.9% term 2 ($p = 0.671$, Fisher's exact test).

I also do not observe major differences in observable characteristics between these groups of vocational students in term 1 and term 2. Tables A.11 and A.12 report balance checks for those apprentices that might have been systematically assigned to different terms. We see that *grade point average* is a significant predictor for treatment assignment, *active in a union* is marginally significant. The data seems, therefore, to be slightly less balanced than we would expect with 30 variables under randomization.

As a further check, I exclude those apprentices from the analysis and re-estimate the effects of schooling with the remaining apprentices. The results suggest that my main finding is robust against potential imbalances that could have been induced by self-selection. Table A.13 provides estimates for the effect of schooling on political interest. All estimates are positive and significant at the 5% level.

Table A.14 shows estimates for the effect of schooling on the likelihood to donate for a specific party. Compared to the large and highly significant effects of schooling on political interest, the effects on the likelihood to support *ÖVP* are already in the full sample smaller and less significant. In my subsample analysis the point estimates are even smaller and not significant anymore.

Table A.11: Balance Table I - Apprentices whose terms were suggested by firms

	Control Groups		Treatment Groups		Test for Difference		
	Mean	SD	Mean	SD	β	95% CI	p-value
Age	17.97	2.24.	18.22	2.06	0.018	[-0.018, 0.058]	0.308
Female	0.35	0.48	0.37	0.49	-0.034	[-0.258, 0.226]	0.744
Austrian citizen	0.93	0.26	0.86	0.35	-0.166	[-0.396, 0.057]	0.120
First apprenticeship	0.95	0.22	0.94	0.24	-0.078	[-0.364, 0.212]	0.550
Work experience†	0.44	1.04	0.59	1.13	0.034	[-0.022, 0.090]	0.206
Income†	5. 53	1.33	5.53	1.51	-0.009	[-0.074, 0.047]	0.738
Exempted from							
civic education	0.21	0.41	0.18	0.38	-0.134	[-0.426, 0.156]	0.219
Attend boarding school	0.29	0.45	0.40	0.49	0.130	[-0.088, 0.361]	0.234
Grade point average	1.87	0.80	2.17	0.85	0.098**	[0.018, 0.169]	0.025
Active in a union†	0.34	0.48	0.47	0.50	0.148*	[-0.002, 0.321]	0.054
Active in a party†	0.07	0.25	0.11	0.31	0.112	[-0.110, 0.351]	0.323
Vocational training							
started in summer	0.81	0.40	0.79	0.41	-0.051	[-0.285, 0.151]	0.603
N	180		147				

†Less observations due to non-responses (Income: $N = 319$, Active in a union: $N = 275$, Active in a party: $N = 288$). Income was measured in steps of €150. *5* denotes net income of €750 to €900 and *6* denotes €900 to €1050, for example.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.12: Balance Table II - Apprentices whose terms were suggested by firms

	Control Groups		Treatment Groups		Test for Difference		
	N	%	*N*	%	β	95% CI	p-value
Company size							**0.576**
0-9 employees	19	10.56	19	12.93	0.083	[-0.109, 0.275]	0.352
10-49 employees	35	19.44	39	26.53	0.116	[-0.120, 0.316]	0.248
50-249 employees	26	14.44	20	13.61	-0.055	[-0.326, 0.284]	0.698
250 employees or more	100	55.56	69	46.94	-0.095	[-0.296, 0.110]	0.282
Hometown							**0.770**
<5.000 inhabitants	112	62.22	83	56.46	-0.051	[-0.179, 0.093]	0.422
5.000 to 20.000	27	15.00	29	19.73	0.081	[-0.041, 0.206]	0.184
20.000 to 100.000	20	11.11	18	12.24	-0.014	[-0.216, 0.170]	0.877
>100.000 inhabitants:							
Center of a city	13	7.22	9	6.12	-0.035	[-0.317, 0.245]	0.776
Outskirts of a city	8	4.44	8	5.44	0.092	[-0.215, 0.388]	0.472
A-levels (Matura)							**0.371**
Have Matura degree	8	4.44	8	5.44	0.097	[-0.225, 0.569]	0.546
Prepare for Matura	45	25.00	28	19.05	-0.062	[-0.176, 0.042]	0.211
Neither have Matura nor prepare for it	127	70.56	111	75.51	0.035	[-0.091, 0.173]	0.558
Country of birth							**0.681**
Austria	169	93.89	131	89.12	-0.029	[-0.114, 0.035]	0.482
Other EU country	4	2.22	8	5.44	0.029	[-0.033, 0.105]	0.458
Non-EU country	7	3.89	8	5.44	0.000	[-0.041, 0.046]	0.996
Parents' country of birth							**0.492**
Both born in Austria	136	75.56	93	63.27	-0.097	[-0.278, 0.069]	0.229
One born in Austria	17	9.44	17	11.56	0.030	[-0.049, 0.106]	0.415
None born in Austria	27	15.00	37	25.17	0.067	[-0.053, 0.196]	0.277
N	180		147				

Notes. For the joint significance tests (p-values in bolt), the dummy for *treatment assignment* was regressed on the categorical variable.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.13: Effects of schooling on political interest. Without self-selection

	(1)	(2)	(3)	(4)
Treated	0.791**	0.812**	0.925**	1.002***
	[0.149, 1.393]	[0.163, 1.411]	[0.245, 1.549]	[0.350, 1.601]
Age		0.006	-0.024	-0.037
		[-0.113, 0.080]	[-0.156, 0.031]	[-0.203, 0.011]
Female		-0.474	-0.569	-0.472
		[-1.270, 0.363]	[-1.317, 0.235]	[-1.236, 0.324]
Grade point average	No	No	Yes	Yes
Additional controls	No	No	No	Yes
N (apprentices)	436	436	436	436
N (class pairs)	17	17	17	17

Notes. Dependent variable is *political interest*, from 0 (not interested at all) to 10 (very interested). 95% confidence intervals in brackets. Additional controls account for *size of training company* and *A-levels*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.14: Schooling and party support. Without self-selection

	(1)	(2)	(3)	(4)
ÖVP	0.015	0.022	0.028	0.031
	[-0.073, 0.100]	[-0.058, 0.099]	[-0.049, 0.102]	[-0.041, 0.102]
SPÖ	0.013	0.015	0.007	0.013
	[-0.062, 0.089]	[-0.058, 0.089]	[-0.065, 0.079]	[-0.057, 0.084]
FPÖ	0.006	0.005	0.001	-0.005
	[-0.079, 0.094]	[-0.074, 0.090]	[-0.085, 0.092]	[-0.090, 0.084]
GREENS	-0.018	-0.020	-0.018	-0.021
	[-0.072, 0.036]	[-0.076, 0.034]	[-0.071, 0.034]	[-0.077, 0.034]
NEOS	0.033	0.035	0.039	0.035
	[-0.017, 0.085]	[-0.014, 0.084]	[-0.011, 0.087]	[-0.015, 0.085]
Other Parties	0.016	0.016	0.016	0.015
	[-0.007, 0.043]	[-0.006, 0.040]	[-0.007, 0.040]	[-0.008, 0.042]
No Party	-0.065	-0.073	-0.073	-0.068
	[-0.185, 0.053]	[-0.193, 0.043]	[-0.196, 0.047]	[-0.196, 0.058]
Included Covariates				
Age and gender	No	Yes	Yes	Yes
Grade point average	No	No	Yes	Yes
Additional controls	No	No	No	Yes
N (apprentices)	436	436	436	436
N (class pairs)	17	17	17	17

Notes. Each entry shows the estimate of the effect of schooling on the likelihood to support a certain political party. 95% confidence intervals in brackets. Additional controls account for *size of training company* and *A-levels*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.C Balance Checks Ibiza

Table A.15: Balance table I - Term 4 as treatment (*Ibiza* affair)

	Control Groups		Ibiza Groups		Test for Difference		
	Mean	SD	Mean	SD	β	95% CI	p-value
Age	18.25	2.51	18.36	2.61	0.283	[-0.169, 0.818]	0.222
Female	0.42	0.49	0.36	0.48	-0.017	[-0.068, 0.033]	0.468
Austrian citizen	0.87	0.33	0.92	0.28	0.050*	[-0.009, 0.103]	0.088
First apprenticeship	0.96	0.20	0.93	0.26	-0.047*	[-0.108, 0.006]	0.083
Work experience†	0.51	1.09	0.54	1.19	0.167	[-0.038, 0.393]	0.106
Income†	5.38	1.28	5.59	1.67	0.102	[-0.215, 0.446]	0.493
Exempted from civic education	0.20	0.40	0.21	0.41	0.029	[-0.036, 0.099]	0.375
Attend boarding school	0.34	0.47	0.29	0.45	0.002	[-0.058, 0.067]	0.934
Grade point average	2.08	0.82	1.88	0.76	-0.189**	[-0.349, -0.032]	0.024
Active in a union†	0.31	0.46	0.33	0.47	-0.029	[-0.152, 0.088]	0.601
Active in a party†	0.07	0.26	0.09	0.29	0.016	[-0.030, 0.065]	0.580
Vocational training started in summer	0.76	0.43	0.71	0.46	-0.089*	[-0.183, 0.004]	0.060
N	563		200				

†Less observations due to non-responses (Income: $N = 741$, Active in a union: $N = 617$, Active in a party: $N = 658$). Income was measured in steps of €150. *5* denotes net income of €750 to €900 and *6* denotes €900 to €1050, for example.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.16: Balance table II - Term 4 as treatment (*Ibiza* affair)

	Control Groups		Ibiza Groups		Test for Difference		
	N	%	*N*	%	β	95% CI	p-value
Company size							**0.788**
0-9 employees	91	16.16	16	8.00	-0.018	[-0.071, 0.036]	0.454
10-49 employees	180	31.97	47	23.50	-0.002	[-0.057, 0.053]	0.937
50-249 employees	79	14.03	44	22.00	0.041	[-0.052, 0.125]	0.322
250 employees or more	213	37.83	93	46.50	-0.021	[-0.114, 0.079]	0.623
Hometown							**0.211**
<5.000 inhabitants	310	55.06	119	59.50	0.052	[-0.021, 0.120]	0.137
5.000 to 20.000	123	21.85	31	15.50	-0.058*	[-0.131, 0.010]	0.091
20.000 to 100.000	62	11.01	22	11.00	-0.013	[-0.074, 0.054]	0.657
>100.000 inhabitants:							
Center of a city	39	6.93	16	8.00	0.010	[-0.034, 0.052]	0.648
Outskirts of a city	29	5.15	12	6.00	0.009	[-0.033, 0.054]	0.682
A-levels (Matura)							**0.789**
Have Matura degree	35	6.22	10	5.00	0.007	[-0.031, 0.046]	0.748
Prepare for Matura	97	17.23	47	23.50	0.021	[-0.044, 0.087]	0.498
Neither have Matura nor prepare for it	431	76.55	143	71.50	-0.027	[-0.108, 0.047]	0.488
N	563		200				

Notes. For the joint significance tests (p-values in bolt), the dummy for *class assignment* was regressed on the categorical variable.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.D Questionnaire

Welcome!

This short survey
is addressed to more than 600
apprentices in Austria.

The survey is conducted
by **Dominik Stelzeneder**,
University of Vienna,
as part of a **research project**.

Your data will be collected **anonymously**
and treated in **strict confidence**.
Participation is **voluntary**; you can leave the survey at any time.

The survey will take about 20 minutes.

Participant information and declaration of consent

I have read the participant information and the declaration of consent. I agree to participate in the survey. ☐

VOTER ELIGIBILITY: 3 ITEMS

Thank you very much for your participation! We begin with two introductory questions on voter eligibility…

Q1. In which year are you born?

Q2. *[If Q1 = 2003]* Are you born before September 29, 2003?

yes 1
no 2

Q3. Are you an Austrian citizen?

yes 1
no 2

INTRODUCTORY QUESTIONS: 4 ITEMS

Thanks! Now we will ask you some content issues.

Q4. Talking very general, are you interested in politics...

Not interested at all 0
to
very interested 10

Q5. Which of the following **political issues** are very important for you currently? You can select up to three issues.
[*RANDOMIZE ITEMS*]

Item 1: Economic Policy (e.g. economic growth)
Item 2: Social Inequality
Item 3: Labor Market and Employment (e.g. unemployment)
Item 4: Immigration
Item 5: Security and Crime
Item 6: Terrorism
Item 7: Climate Change
Item 8: Party financing (e.g. regulation of campaign contributions, transparency)
Item 9: European Integration (e.g. European cohesion)
Item 10: Health and Nursing Care
Item 11: Corruption

Item 13: Pensions and Old-Age Provision
Item 13: Tax Policy
Item 14: Asylum Policy
Item 15: Education and Research
Item 16: Housing
Item 17: Environmental Protection
Item 18: Judicial System
Item 19: Media (e.g. regulation, press freedom, fake news)
Item 20: No Issue
Item 21: Other Issue [*please specify:*]

Q6. In politics people often talk of "left" and "right". Where would you place the following parties on a scale from 0 to 10, where 0 means "left" and 10 means "right"? *[RANDOMIZE ITEMS]*

Item 1: ÖVP
Item 2: SPÖ
Item 3: FPÖ
Item 4: NEOS
Item 5: GREENS

left..... 0
to
right..... 10

I don't know this party 77
Don't know..... 88

Q7. Where would you place yourself on this scale from 0 to 10, where 0 means "left" and 10 means "right"?

left..... 0
to
right..... 10

Don't know..... 88
Refused..... 99

POLITICAL PARTICIPATION and VOTING BEHAVIOR: 8 ITEMS

The following part of the survey is about participation in elections and voting decisions.

Q8. *[if Q1 < 2003 and Q3 = 1 or if Q1 = 2003 and Q2 = 1 und Q3 = 1]* The last parliamentary election was on September 29, 2019. If the election would be next Sunday instead, how likely would you turn out to vote?

Would certainly not turn out to vote 1
to
would certainly turn out to vote 10

Q9. *[if Q1 > 2003 or Q3 = 2 or if Q1 = 2003 and Q2 = 2]* The last parliamentary election was on September 29, 2019. If the election would be next Sunday instead and you were entitled to vote, how likely would you turn out to vote?

Would certainly not turn out to vote 1
to
would certainly turn out to vote 10

Q10. Now suppose, the next parliamentary election would indeed be held next Sunday. Which party would you vote for?

ÖVP 1
SPÖ 2
FPÖ 3
NEOS 4
JETZT 5
GREENS 6
KPÖ 7
WANDL 8
Invalid vote……..... 66
Don't know 88
Refused 99

Q11. At the last parliamentary election on September 29, 2019, many citizens did not go to the election for good reason. How about you? Which statement applies to you?

I was not entitled to vote, therefore I did not turn out to vote .....……..... 0
I did not turn out to vote at the parliamentary election on September 29, 2019 1
I thought about turning out to vote but I did not this time 2
I am sure that I turned out to vote on September 29, 2019 3

Q12. *[if Q11 = 3]* Which party did you vote for in the last national election on September 29, 2019?

ÖVP 1
SPÖ 2
FPÖ 3
NEOS 4
JETZT 5
GREENS 6
KPÖ 7
WANDL 8
Invalid vote 66
Refused …..... 99

Q13. *[if Q11 ≠ 3]* Suppose you would have turned out to vote in the national election on September 29, 2019. Which party would you have voted for?

ÖVP 1
SPÖ 2
FPÖ 3
NEOS 4
JETZT 5
GREENS 6
KPÖ 7
WANDL 8
Invalid vote 66
Don't know 88
Refused …..... 99

Q14. On a scale from 0 to 10, how likely is it that you will ever vote for the following parties? 0 means „highly unlikely" and 10 means „very likely".

Item 1: ÖVP
Item 2: SPÖ
Item 3: FPÖ
Item 4: GREENS
Item 5: NEOS

Highly unlikely 0
to
very likely 10

I don't know this party 77
Don't know 88

Q15. For your participation in this survey, we will donate money to a political party. Every participant can donate € 10. Which party would you like us to support with a donation of € 10?*

ÖVP 1
SPÖ 2
FPÖ 3
NEOS 4
JETZT 5
GREENS 6
KPÖ 7
WANDL 8
No party 12

*In total, more than 600 apprentices take part in this study. After every apprentice has made a choice, the computer will randomly select 20% of the choices and we will donate the money based on those randomly selected choices. To prove that we indeed donated the money to the parties, we will publish scanned receipts of the transactions and publish them for three months, starting on December 1, 2019, on https://homepage.univie.ac.at/dominik.stelzeneder/

The donations are legal and they support the parties doing their work.

ATTITUDES TOWARDS DEMOCRACY IN AUSTRIA: 6 ITEMS

Now we want to ask you some general questions about Austrian politics and about political issues.

Q16. How satisfied or unsatisfied are you in general with how democracy works in Austria? Are you...

Very unsatisfied 0
to
very satisfied 10

Q17. Please indicate how strongly you agree or disagree with the following statements.
[RANDOMIZE ITEMS]

Item 1: "I have a good understanding of the most important political topics in Austria."
Item 2: "I am confident with actively participating in political discussions."
Item 3: "Politicians do not care about what people like me think."
Item 4: "No matter who I vote for, it would not change anything in Austrian politics."

I fully agree..... 0
to
I fully disagree..... 10

Q18. When you think about Austria's future, how optimistic or pessimistic are you?

very optimistic..... 0
to
very pessimistic..... 10

Q19. Please indicate how strongly you agree or disagree with the following statements.
[RANDOMIZE ITEMS]

Item 1: "Income differences are too large in Austria."
Item 2: "The welfare state makes people languid and lazy."
Item 3: "The police needs more competences to enforce law and order."
Item 4: "Immigrants are taking the jobs of Austrians'."

I fully agree..... 0
to
I fully disagree..... 10

Q20. Please indicate how strongly you agree or disagree with the following statements.
[RANDOMIZE ITEMS]

Item 1: "When talking about 'compromises' in politics, one actually means that one is betraying one's principles."
Item 2: "Most politicians only care about the interests of the rich and powerful."
Item 3: "Most politicians are trustworthy."
Item 4: "The parties are the major problem in Austria."

I fully agree..... 0
to
I fully disagree..... 10

Q21. Please indicate how strongly you agree or disagree with the following statements.
[RANDOMIZE ITEMS]

Item 1: "A strong leader in government is good for Austria, even if he/she does not always follow the rules to move things forward."
Item 2: "The people should make important political decisions, not politicians."
Item 3: „Corporations and not the government decide over politics."

I fully agree..... 0
to
I fully disagree..... 10

„VOTING WITH 16": 3 ITEMS

Next, we will ask you three questions about conversations and discussions about politics.

Q22. How often have you talked about politics with these people or groups?
[RANDOMIZE ITEMS]

Item 1 Father
Item 2 Mother
Item 3 Workmates
Item 4 Teachers
Item 5 Fellow students
Item 6 Other friends
Item 7 Siblings

daily..... 1
almost daily..... 2
several times per week..... 3

several times per month 4
hardly ever..... 5
never..... 6

not applicable 77

Q23. How interested are these people or groups in politics?
[RANDOMIZE ITEMS]

Item 1 Father
Item 2 Mother
Item 3 Workmates
Item 4 Teachers
Item 5 Fellow students
Item 6 Other friends
Item 7 Siblings

Very interested 0
to
not at all interested 10

not applicable 77
don't know 88

Q24. In the current term in school, have you…
[RANDOMIZE ITEMS]

Item 1: discussed about Austrian politics?
Item 2: talked about the parliamentary election 2019?
Item 3: participated in school democracy, e.g. have you participated in the election of the school representative or have you participated in a student assembly?
Item 4: participated in a project that was related to politics?

yes 1
no..... 2
don't know 88

POLITICAL ATTITUDES: 13 ITEMS

Now we focus on political topics. What do you think about the following questions?

Q25. Should the weekly working time be reduced to 30 hours without wage adjustment?
Q26. Should company taxes[1] be reduced?
Q27. Should bequests remain tax-exempt?
Q28. Should the legal retirement age be raised?
Q29. Should an independent agency be established that investigates police violence?
Q30. Should the joint school for 10 to 14 year-olds be realized?
Q31. Should asylum seekers get a work permit?
Q32. Should criminal foreigners be deported, independent of their duration of stay in Austria?
Q33. Should the general right to vote in national elections remain restricted to Austrian citizens?
Q34. Should the Court of Auditors be allowed to audit all party finances and to impose penalties?
Q35. Should the compulsory broadcasting fees („GIS Gebühren“) be abolished?
Q36. Should internet forums be obliged to record the real names[2] of all its users and to release them if necessary?
Q37. Should a CO2-tax[3] be introduced?

yes 1
no 2

no answer 88

AND each time: How important is this topic to you?

unimportant 1
to
important 9

[1] Company taxes refer to all taxes that companies have to pay. These taxes include profit taxes (like income tax, corporate income tax, and business tax), consumption taxation (value added tax and real estate transfer tax), and property taxation (land tax, inheritance tax and gift tax).
[2] Real name (“*Klarname”)* means the real and official first and second name, in contrast to nicknames or pseudonyms of users in social networks or internet forums.
[3] A CO2-tax taxes carbon dioxide and other greenhouse gases that substantially contribute to the climate change crisis. The subject of taxation is mainly the combustion of fossil energy sources, for example in the industry, but also fuels.

KNOWLEDGE QUESTIONS ABOUT PARTY POSITIONS: 5 ITEMS

Q38. And what do the parties think about the following question:

Should the weekly working time be reduced to 30 hours without wage adjustment?

Q39. And what do the parties think about the following question:

Should company taxes[4] be reduced?

Q40. And what do the parties think about the following question:

Should the joint school for 10 to 14 year-olds be realized?

Q41. And what do the parties think about the following question:

Should criminal foreigners be deported, independent of their duration of stay in Austria?

Q42. And what do the parties think about the following question:

Should a CO2-tax[5] be introduced?

Item 1: ÖVP
Item 2: SPÖ
Item 3: FPÖ
Item 4: GREENS
Item 5: NEOS

I am sure: the party agrees 1
I believe: this party agrees 2
I don't know 3
I believe: the party disagrees 4
I am sure: the party disagrees 5

[4] Company taxes refer to all taxes that companies have to pay. These taxes include profit taxes (like income tax, corporate income tax, and business tax), consumption taxation (value added tax and real estate transfer tax), and property taxation (land tax, inheritance tax and gift tax).
[5] A CO2-tax taxes carbon dioxide and other greenhouse gases that substantially contribute to the climate change crisis. The subject of taxation is mainly the combustion of fossil energy sources, for example in the industry, but also fuels.

SOCIO-DEMOGRAPHIC FACTORS and BACKGROUND: 26 ITEMS

Finally, some questions for statistical purposes.

SD1. Gender

male 1
female 2

SD2. In which month has your apprenticeship begun?

SD3. How many years had you worked in a company before your apprenticeship begun?[6]

SD4. Have you already completed another apprenticeship?

SD5. Are you part of the program „**Lehre mit Matura**[7]“?

yes 1
no 2
no, I already have the general qualification for university entrance 3

SD6. Do you get exactly the legally required minimum remuneration or do you get more than the legally required minimum remuneration?

I get exactly the legally required minimum remuneration 1
I get more than the legally required minimum remuneration..... 2

SD7. What is your net salary after taxation and social security contributions?

Less than 300€ 1
300 to 450€ 2
450 to 600€..... 3
600 to 750€..... 4

[6] If you had not worked in a company before, please select „0 years“
[7] This program combines an apprenticeship with additional theoretical courses to acquire the general qualification for university entrance.

750 to 900€ 5
900 to 1050€ 6
1050 to 1200€ 7
1200 to 1350€ 8
1350 to 1500€ 9
1500 to 1650€ 10
1650 to 1800€ 11
1800€ or more 12

No answer 99

SD8. How many employees does your company have in total?

0 – 9 1
10 – 49 2
50 – 249 3
250 or more 4

SD9. Did your employer ask you about your preferred term of schooling this year?

Yes and, therefore, I attend the current school term 1
Yes, however, I preferred a different term 2
No, my company has decided 3
No, the vocational school has decided 4
Don't know 5

SD10. In which term did you attend school last year?

In the first term (September-November) 1
In the second term (November-January) 2
In the third term (February-April) 3
In the fourth term (April-June) 4

SD11. Are you exempted from the subject „Civic Education"?

yes 1
no 2

SD12. Do you stay in a boarding school?

yes 1
no 2

SD13. With whom do you usually live in a household?

Father 1
Mother 2
Siblings 3
With one or more friends 4
With one or more friends 5
Husband, wife oder partner 6
It's only me 7
Others: *[please specify:]* 8

*Please check all that apply.

SD14. Which grade did you get on average last school year?

SD15. Have you ever visited the website "wahlkabine.at"?

yes 1
no 2

I don't know this website 99

SD16. *[if SD15 = 1]* On "wahlkabine.at", have you…

Item 1 answered the questions on the national election 2019?
Item 2 enquired about party positions?

yes 1
no 2

SD17. *[if SD15 = 1]* Who has told you about the website „wahlkabine.at“?

Item 1 Parents
Item 2 Workmates
Item 3 Teachers
Item 4 Fellow students
Item 5 Other friends
Item 6 Others: *[please specify]*

*Please check all that apply.

SD18. In which country are you born?

Austria 1
In another country in the EU 2
In a country outside of the EU 3

SD19. Are your parents born in Austria?
[NOTE: It’s about your biological parents]

Yes, both of them 1
Yes, one of them 2
No, none of them 3

SD20. What is the educational attainment of your father?

Secondary school 4
Vocational school 7
Vocational middle school [e.g. HASCH] 8
High school with Matura [e.g. HTL, HAK, HBLA] 10
University 13
Other [*please specify*] 16

Does not apply 77
Don’t know 88

SD21. What is the educational attainment of your mother?

Secondary school 4
Vocational school 7
Vocational middle school [e.g. HASCH] 8
High school with Matura [e.g. HTL, HAK, HBLA] 10
University 13
Other [*please specify*] 16

Does not apply 77
Don't know 88

SD22. Are you member of a union?

yes 1
no 2

no answer 3

SD23. Are you active in a political association or a political organization?

yes 1
no 2

no answer 3

SD24. Where do you live?

Village [less than 5.000 inhabitants] 1
Small town [5.000 to 20.000 inhabitants] 2
Town [20.000 to 100.000 inhabitants] 3
In the center of a city [more than 100.000 inhabitants] 4
On the outskirts of a city 5

SD25. How many siblings do you have?

0 0
1 1
2 2
3 3
4 4
More than 4 5

SD26. And how many of your siblings are older than you?

0 0
1 1
2 2
3 3
4 4
More than 4 5

END: 1 Item

Q43. Do you want to tell us something about the survey or individual questions?

[OPEN ANSWER]: _____

Thank you very much for your participation in the survey!

Please remain seated; the other participants will also finish soon.

Chapter 2

Gender Differences in Competitiveness: Does the Framing Matter? Experimental Evidence[1]

Previous studies show that men are more competitive than women in stereotypically male tasks but not in stereotypically female tasks. However, to study the effect of a task's domain, these studies change the entire task and not the domain only. In our experiment we change the perceived domain of the task without changing the task itself: Participants solved *Raven's Progressive Matrices* that were either framed as *intelligence* tasks or as *creativity* tasks but strictly identical otherwise. This treatment affected the perceived gender stereotype of the task and both men's and women's competitive performance. The gender difference in competitiveness was, however, not affected by our treatment.

2.1 Introduction

In this paper we study whether the domain of a task matters for men's and women's competitiveness. Gneezy, Niederle, and Rustichini (2003) showed that women performed poorer in competitions than men - even though their performance did not differ in a noncompetitive setting. Moreover, women seem to shy away from competition. In their seminal study, Niederle and Vesterlund (2007) showed that women are only about half as likely as men to enter a competition when given a noncompetitive alternative. Factors like differences in task performance, self-confidence, and risk aversion were not able to explain the gap between men's and women's tournament entry rates. This finding triggered a large and very fruitful academic debate about the causes and consequences of these gender differences in competitiveness (see Niederle and Vesterlund 2011 and Niederle 2016 for

[1]This chapter is joint work with Kai Barron and Sebastian Schweighofer-Kodritsch.

reviews of the literature).

Ensuing papers typically use small modifications of the original Niederle-Vesterlund design as treatments to test whether the modifications affect the gender gap or not. In this way several factors that influence the size (or even the direction) of the gender gap in competitiveness have been established. The domain of the task that is used in experiments has been shown to be of particular importance. Numerous studies used two or more different tasks to investigate whether men's and women's competitiveness depends on the domain of the task. If the task was perceived as stereotypically "female" or "neutral", and *not* as stereotypically "male", the gender gaps in competitive performance and tournament entry rates were either smaller or even reversed (Günther et al. 2010, Dreber, Essen, and Ranehill 2011, Cárdenas et al. 2012, Shurchkov 2012, Wieland and Sarin 2012, Grosse, Riener, and Dertwinkel-Kalt 2014, and Iriberri and Rey-Biel 2017).[2]

However, all these papers study the effect of varying the domain of a task by changing the *entire* task but not by changing the domain of the task only. In our experiment we use the same task for every participant and we only vary how the task is framed. We use *Raven's Progressive Matrices* (Raven, Raven, and Court 1998) and refer to them as either *creativity tasks* or *intelligence tasks*. Consequently, we only change the *domain* of the task and not the task itself. This design rules out spurious results that could arise from correlations between the domain of a task and other task specific characteristics that might affect the competitiveness of men or women.

For a policy maker it is highly relevant whether the domain of a task per se causally matters or not. If it does not, then dismantling gender stereotypes will not affect gender differences in competitiveness. Furthermore, if it matters how tasks are framed, gender equality in competitive settings could be increased very easily. Institutions that perform competitive entrance exams could adjust the frames of examination tasks to simultaneously allot available places to the best performing applicants and to improve gender balance of (successful) applicants.

In our experiment, we investigate whether and how the domain of the task affects competitiveness. The *creativity* and the *intelligence* domains are used in our study for two reasons. First, we consider both labels credible for the *Raven Matrices*. Second, we supposed that gender stereotypes differ between both domains. We conjecture that the *Raven Matrices* are perceived as more stereotypically "female" under the *creativity* frame than under the *intelligence* frame. Based on the previous literature, we hypothesize that this difference in the perceived gender stereotype of the task affects men's and

[2]Furthermore, Hernandez-Arenaz (2020) observes that the perceived "maleness" of a task correlates with men's self-confidence and with men's likelihood to enter a competition. Bordalo et al. (2019) complements this observation by showing that gender stereotypes across different domains contribute to gender differences in self-confidence and in assessments of others.

women's competitiveness differently. We hypothesize that women's performance and their willingness to enter a competition are higher under the *creativity* frame than under the *intelligence* frame and we expect the opposite for men.

What could be reasons for differences in competitiveness depending on the domain of a task? The recent literature has mainly focused on the role of gender stereotypes to answer this question. There are two broad channels through which a stereotype can affect competitiveness. First, it can influence one's self-assessment by inducing a bias towards the stereotype of one's own gender (Wieland and Sarin 2012 and Hernandez-Arenaz 2020). Lower perceived chances of winning a tournament in turn might lead to lower effort and to lower tournament entry rates. Second, a gender stereotype can affect competitiveness via inducing *stereotype threat*, which is described as "[...]being at risk of confirming, as self-characteristic, a negative stereotype about one's group" (Steele and Aronson 1995, p. 797). In contrast to the first channel, the individual now does not need to believe in the stereotype, it is sufficient that he or she thinks that others share the stereotype. Besides stereotypes, ability differences that vary with the domain of a task are also discussed in the literature as a reason for different gender gaps in competitiveness. Women seem to be better at tasks that demand attention to detail. In a competition, they thus concentrate more on the quality of work output while men focus more on the quantity of work output (Shurchkov 2012). Depending on the specific task, men or women are hence better off. Another potential reason for varying gender gaps in competitiveness are differences in utility functions. Men and women possibly derive more or less utility from being competitive in different domains: They might consider *winning a tournament* as more or less important depending on the domain of the task. Already Atkinson (1958) shows that competitive performance increases when the incentive to win is higher (in his study this was achieved by increasing the prize money).

Our results indeed show that the perceived gender stereotype of the tasks were affected by our treatment. Furthermore, men and women performed significantly better in the tournament when the tasks were framed as *intelligence* tasks and not as *creativity* tasks. However, we do not find evidence for *differences* in framing effects on men's and women's competitiveness: Men and women performed equally well in the tournament, and decisions to either enter a competition or to select a non-competitive alternative were not affected. The exogenous change in the gender stereotype of the tasks did thus not affect men's or women's competitiveness differently. Effects of gender differences in domain specific abilities are ruled out by our design since the tasks are strictly identical for all participants in our study. We, therefore, conjecture that men and women were both more competitive in the *intelligence* frame because they derive more utility from performing well in competitions in an *intelligence* domain than in a *creativity* domain.

2.2 Experimental Design

Our experiment was based on the design of van Veldhuizen (2022). It had five tasks and at the end of the experiment one of them was randomly selected for payment. The full instructions are provided in Appendix 2.A.

The first three tasks of the experiment followed the original design of Niederle and Vesterlund (2007). In each of these three tasks, participants had 2 minutes to solve a series of *Raven's Progressive Matrices* (Raven, Raven, and Court 1998). The matrices contained either six or nine separate figures that are related to one another in a particular way. In each matrix, one of the nine figures was missing. The participants had to select the correct figure that completed the matrix. After submitting an answer, participants learned whether their answer was correct and they were informed about the total number of matrices that they had solved correctly in the task. Before Task 1 started, we showed them a practice exercise to make them familiar with the task.

In Task 1 (*piece rate*) participants received \$0.25 for each matrix they solved correctly. In Task 2 (*tournament*) they were randomly assigned to groups of four. In this task, only the best performing participant of each group received a payment. The winners were paid \$1.00 for each correct solution and the others received \$0 (in case of a tie, the winner was randomly determined). In Task 3 (*decision*) participants had to choose between applying the *piece rate* or the *tournament* payment scheme for their next performance. If they chose the *tournament* payment scheme, they competed against their group members' performance in Task 2. If they solved more matrices in Task 3 than their group members did in Task 2, they received \$1 for each correct solution and \$0 otherwise. Decisions in Task 3 did hence not affect other participants' decisions or earnings. If Task 3 was randomly selected for payment we informed participants about whether they would have won the tournament or not, independent of the payment scheme they chose.[3]

Tasks 4 and 5 were first added to the experimental design of Niederle and Vesterlund (2007) by van Veldhuizen (2022). In Task 4 we elicited participants' beliefs about how well they performed in Task 2 relative to their group members. They were asked to answer the following question: "How likely do you think it is that you won the tournament in Task 2?".

After the participant had entered a number, the computer randomly chose another number between 0 and 100. If this randomly chosen number was larger than the number that was entered, then the participant received \$4 with probability equal to the value of the randomly chosen number (in percentage points). If the randomly chosen number was smaller than or equal to the number that was entered, then the participant received \$4

[3]This was a small modification that we implemented to rule out effects of feedback aversion.

if he or she had won the tournament in Task 2 and \$0 otherwise.

This procedure made sure that it was incentive compatible for a utility maximizing subject with any risk preferences to state his or her true belief (Karni 2009). Since we did not use a subject pool of university students, the instructions for this task were slightly shortened and simplified compared to the original version of van Veldhuizen (2022).

In Task 5 subjects had to make a series of 20 decisions between a safe *Option A* and a risky *Option B*. The payment sizes were determined by the subject i's number of correctly solved matrices in Task 2 (tournament), x_i, and they were the same in all 20 decision problems. *Option A* was a sure payment of $\$0.25x_i$ and *Option B* was a risky lottery that paid either \$0 or $\$x_i$. The 20 decision problems only differed in terms of chance to win the lottery when *Option B* was chosen. In decision problem 1, the chance of winning $\$x_i$ in *Option B* was 5%. This winning chance increased in any of the following decision problems by 5 percentage points, such that in decision problem 20, the chance of winning $\$x_i$ in *Option B* was 100%.

At the end of the experiment the study participants had to answer four questionnaires. We asked them about their age and gender, whether they have seen or even solved the kind of task of Tasks 1 to 3 before, and how much they liked working on these tasks. In addition, they had to answer two incentivized questions: First, they had to state their belief on whether men or women performed better in Task 2 (tournament). Second, we asked about their belief on whether most people think that men or women were better at solving tasks like those in Tasks 1 to 3. Participants received \$0.25 for each correct answer.[4] Furthermore, we elicited two measures of subjects' risk aversion (Eckel and Grossman 2002 and Holt and Laury 2002).

For all participants the experiment was strictly identical, except for the framing of the tasks. They saw the same instructions, they worked on the very same tasks and they were paid according to the same incentive schemes. For 233 participants, however, the tasks were framed as *"intelligence"* tasks, whereas for the other 226 participants the tasks were framed as *"creativity"* tasks. The treatments only differed in these three aspects from each other:

1. Whenever we referred to *Task 1*, *Task 2* or *Task 3*, we included either the term *"intelligence"* or the term *"creativity"* before.

2. The font of the headlines in Tasks 1 to 3 differed. An italic font was used in the *"creativity"* frame, whereas a bold font was used in the *"intelligence"* frame.

3. A small image was added to the headlines of Task 1 to Task 3, a colorful tree

[4] As in Krupka and Weber (2013), the answer of the majority of participants on the second question determined the correct answer to that question.

in the *"creativity"* frame and the shape of a head with gear wheels inside in the *"intelligence"* frame.

Upon completion of the experiment, the participants were informed about their payoffs. They received a show-up fee of \$4, their individual payment from one randomly selected task and a payment depending on their decisions in the questionnaires. Information on payoffs that depended on the performance or on answers of other study participants was sent to the participants after the entire experiment had ended. The average time needed to complete the experiment was 16 minutes and the average earning was \$ 7.60.

We conducted the experiment online. 459 participants (235 women and 224 men) were recruited via *Prolific* and randomly assigned to a treatment. The demographic distribution of our participants is representative for the US population in terms of age, sex, and ethnicity. The experiment was programmed in *oTree* (Chen, Schonger, and Wickens 2016) and it was conducted in November 2021.

2.3 Results

Table 2.1 provides descriptive statistics of the main variables of interest by gender and frame. Before we turn to the effects of our treatment, we first disregard the framing and show results for the pooled sample. The p-values in this section are calculated using two-sided *Fisher's exact tests* for binary outcome variables and two-sided t-tests for continuous outcome variables, if not otherwise indicated.

Table 2.1: Descriptive statistics by gender and frame

	Women			Men		
	(1) Creativity	(2) Intelligence	(3) Pooled	(1) Creativity	(2) Intelligence	(3) Pooled
Score Task 1 (Piece Rate)	3.34	3.43	3.38	3.54	3.43	3.48
	(1.44)	(1.54)	(1.48)	(2.07)	(1.50)	(1.80)
Score Task 2 (Tournament)	4.75	5.27	5.00	4.86	5.27	5.06
	(1.70)	(2.26)	(2.01)	(2.27)	(2.06)	(2.17)
Performance improvement	1.41	1.84	1.62	1.32	1.84	1.58
	(1.51)	(1.72)	(1.62)	(1.94)	(1.65)	(1.81)
Guessed chance of winning Task 2	0.38	0.42	0.40	0.36	0.42	0.39
	(0.27)	(0.29)	(0.28)	(0.27)	(0.27)	(0.27)
Tournament entry	0.33	0.36	0.34	0.43	0.41	0.42
	(0.47)	(0.48)	0.48	(0.50)	(0.49)	(0.49)
Belief that women performed better in Task2	0.62	0.58	0.60	0.62	0.52	0.57
	(0.49)	(0.50)	(0.49)	(0.49)	(0.50)	(0.50)
Belief that most others think that women performed better	0.44	0.22	0.33	0.40	0.29	0.34
	(0.50)	(0.42)	(0.47)	(0.49)	(0.45)	(0.48)
Number of participants	121	114	235	112	112	224

Notes. Averages with standard deviations in parentheses.

2.3.1 Pooled Sample Results

Using the pooled sample that disregards the framing leads to results that are very much in line with the literature. The performance of men and women did not differ in Task 1 (women: 3.38 vs. men: 3.48, $p = 0.519$) or in Task 2 (women: 5.00 vs. men: 5.06, $p = 0.765$). Men in our sample were less risk avers than women. The difference was significant according to the measures of Eckel and Grossman (2002) ($p = 0.001$) and of Task 5 ($p = 0.000$), and in the same direction but not significant according to the measure of Holt and Laury (2002) ($p = 0.410$). To our surprise, there was no significant gender difference between men's and women's guessed chance of having won the tournament in Task 2 (women: 0.40 vs. men: 0.39, $p = 0.888$).

As observed in previous research, men were more likely to enter the competition than women. This gender difference is, however, small compared to results in other papers and only close to the boundary of marginal significance (women: 0.34 vs. men: 0.42, $p = 0.103$). Furthermore, controlling for risk aversion decreases this gender gap even further. In a regression that uses the point in Task 5 at which a participant switches from the risky lottery to the safe option, the gender gap shrinks to a 3.1 percentage point difference ($p = 0.888$).[5]

2.3.2 Treatment Effects

Beliefs about winning the tournament in Task 2 and actual performance in Task 2 were significantly affected by the framing. Figures 2.1 and 2.2 show CDF plots of men's and women's performance in Task 1 and Task 2 by frame. There was no difference in Task 1 performance between participants in the *creativity* frame and in the *intelligence* frame (creativity: 3.43 vs. intelligence: 3.43, $p = 0.978$). Neither the performance of men (creativity: 3.53 vs. intelligence: 3.43, $p = 0.657$) nor the performance of women (creativity: 3.34 vs. intelligence: 3.43, $p = 0.640$) was affected by the frames. Strikingly, this is not the case in Task 2: Participants in the *intelligence* frame performed significantly better than those in the *creativity* frame (creativity: 4.80 vs. intelligence: 5.27, $p = 0.016$). This result is caused by differences in performance improvement from Task 1 to Task 2. Under both frames, participants solved more matrices in Task 2 than in Task 1. But their performance increase was significantly higher in the *intelligence* frame than in the *creativity* frame (women: +1.41 from Task 1 to Task 2 in the *creativity* frame compared to +1.84 in the *intelligence* frame ($p = 0.043$); men: +1.32 in the *creativity* frame compared to +1.84 in the *intelligence* frame ($p = 0.032$)).

[5]If the measure of Holt-Laury (Eckel-Grossman) is used, the gender gap decreases to a 7.5 (5.5) percentage point difference ($p = 0.129$ and $p = 0.228$, respectively).

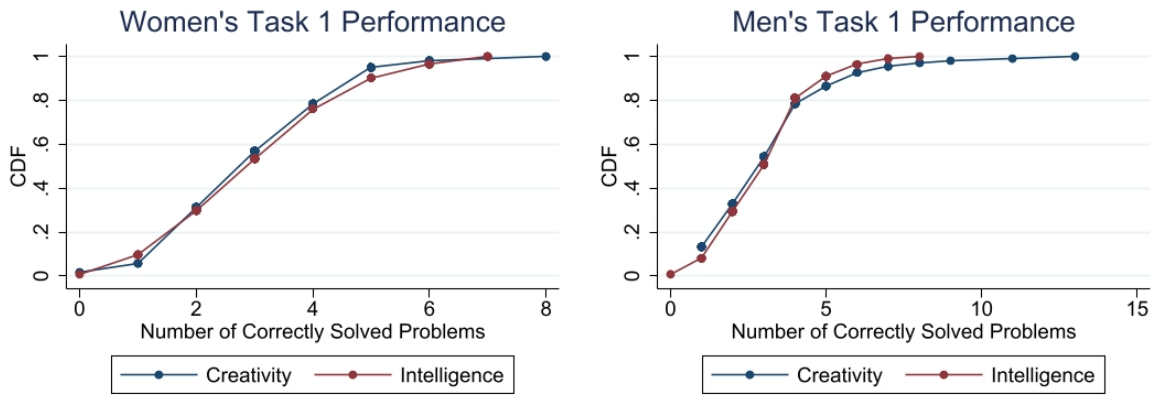


Figure 2.1: CDF of number of correctly solved problems in Task 1 (Piece Rate). Left panel: women; right panel: men

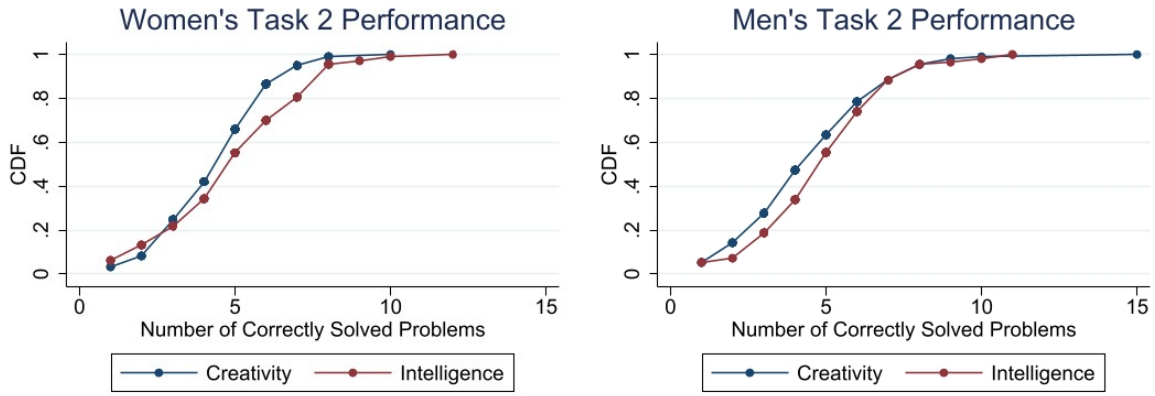


Figure 2.2: CDF of number of correctly solved problems in Task 2 (Tournament). Left panel: women; right panel: men

The differences in performance improvements were accompanied by changes in beliefs about one's likelihood of having won the tournament in Task 2. Participants on average stated a likelihood of 36.9% in the *creativity* frame and 42.0% in the *intelligence* frame ($p = 0.043$). Men's self-confidence (creativity: 36% vs. intelligence: 42%, $p = 0.089$) differed slightly more between both frames than women's (creativity: 38% vs. intelligence: 42%, $p = 0.239$).

The gender gap in tournament entry rates did not change significantly due to the framing of the tasks (creativity: 0.10 vs. intelligence: 0.05, $p = 0.606$, difference-in-difference test). Neither men's (creativity: 0.43 vs. intelligence: 0.41, $p = 0.788$) nor women's (creativity: 0.33 vs. intelligence: 0.36, $p = 0.641$) decisions in Task 3 were affected.

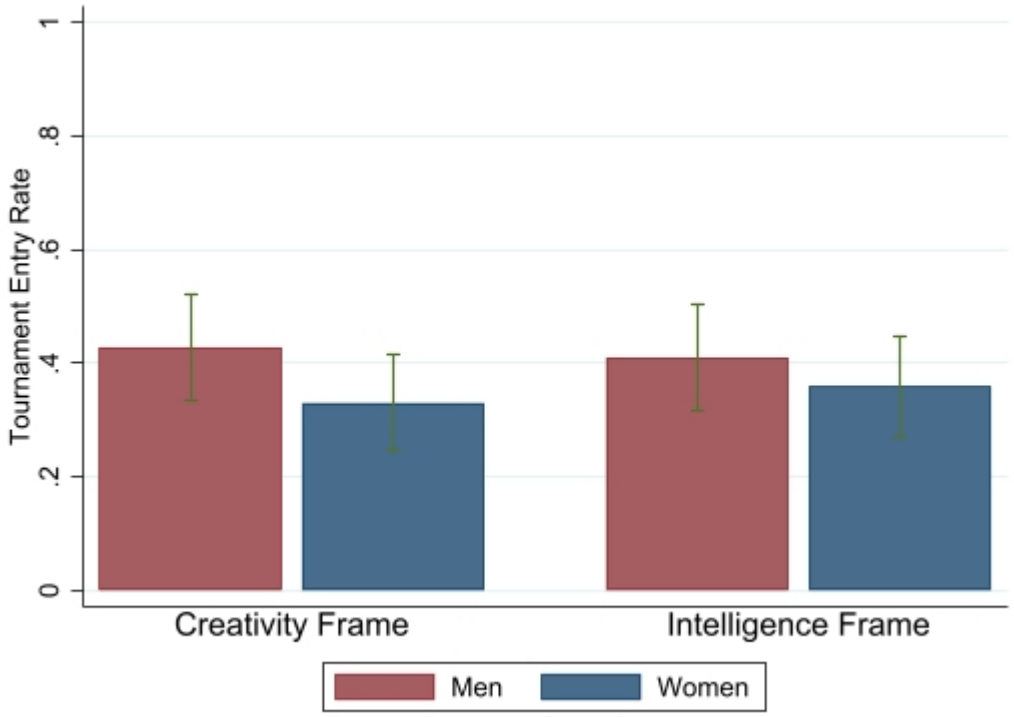


Figure 2.3: Share of participants who chose the tournament payment scheme in Task 3 by frame and gender. The error bars represent 95% confidence intervals.

2.3.3 Frame and Gender Stereotypes

We used the *creativity* and the *intelligence* frames because we considered both labels suitable for the *Raven Matrices* but also because we expected them to have different gender connotations. Answers on the incentivized questions in our questionnaire partly confirm this expectation.

Table 2.2 shows the share of participants who believed that women perform better in Task 2 on average. We see that the framing did not significantly affect this belief. Both more men and more women believed that the task favors women under the *creativity* frame (62% each). Under the *intelligence* frame, slightly (though insignificantly) fewer

Table 2.2: Share of participants who believed that women perform better in Task 2 on average

Frame	Women	Difference to 50%	Men	Difference to 50%
Creativity	62%	p=0.011	62%	p=0.018
Intelligence	58%	p=0.111	52%	p=0.777
Treatment effect	p=0.525		p=0.139	

Notes. p-values are based on t-tests (*treatment effect*) and binomial probability tests (*difference to gender neutral share of 50%*).

Table 2.3: Share of participants who believed that most others think that women perform better

Frame	Women	Difference to 50%	Men	Difference to 50%
Creativity	44%	p=0.203	40%	p=0.047
Intelligence	22%	p=0.000	29%	p=0.000
Treatment effect	p=0.000		p=0.068	

Notes. p-values are based on t-tests (*treatment effect*) and binomial probability tests (*difference to gender neutral share of 50%*).

men and women considered the task as favoring women (men: 52% and women: 58%).

Results for the beliefs about whether most other people think that men or women perform better are in stark contrast to the results above (see Table 2.3). Under the *creativity* frame, 44% of women and 40% of men believed that most others think that women perform better. Under the *intelligence* frame, these shares drop significantly to just 22% and 29%, respectively.

Our frames thus did not affect participants' own beliefs about the gender stereotype of the task but it had a significant effect on the belief about how others perceive the task.

2.3.4 Does the Stereotype Matter?

Previous research has suggested that task specific gender stereotypes might matter for gender differences in competitiveness (Günther et al. 2010, Dreber, Essen, and Ranehill 2011, Shurchkov 2012, Wieland and Sarin 2012, Grosse, Riener, and Dertwinkel-Kalt 2014, Iriberri and Rey-Biel 2017, Hernandez-Arenaz 2020). However, our results do not provide evidence for an effect of the task's gender stereotype on competitiveness.

Table 2.4 shows whether and how participants' confidence in their tournament performance, their actual tournament performance, and their decisions to enter a competition interrelate with the perceived gender stereotype of the task. A significant relationship would be a first indicator for an effect of gender stereotypes on competitiveness (see also Hernandez-Arenaz 2020). In our analyses we consider one's own perception of the task and the belief about how most others perceive the tasks.

We observe two significant relationships.

(1) Women who believed that women perform better on average than men in Task 2 performed better than other women ($p = 0.013$).

(2) Men who thought that most participants expected women to perform better than men were less likely to enter the competition in Task 3 ($p = 0.073$).

Table 2.4: Relationship between perceived task stereotype and competitiveness

	Guessed chance of winning Task 2		Task 2 performance		Tournament entry	
	(1)	(2)	(3)	(4)	(5)	(6)
	Women	Men	Women	Men	Women	Men
Belief that women performed better	0.043	-0.041	0.66**	-0.07	-0.099	-0.096
	(0.037)	(0.036)	(0.26)	(0.29)	(0.063)	(0.067)
Belief that most others think that women performed better	-0.004	-0.020	-0.06	-0.21	-0.094	-0.125*
	(0.038)	(0.038)	(0.28)	(0.31)	(0.376)	(0.069)
Observations	235	224	235	224	235	224

Notes. Each entry shows the effect of either one's own belief or one's belief about the stereotypes of others on the respective column variable. E.g., a woman who beliefs that women performed better scored better in Task 2 by 0.66 correct matrices on average. Two-sided t-tests, standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$

Based on these estimation results, one might conclude that the perceived stereotype of the task would have an effect on these two outcome variables. However, in view of the observed framing effects this does not seem to be the case. We can show that stereotypes were presumably not the main drivers for our observed treatment effects. To do so, we exploit that our frames exogenously affected men's belief about how others perceive the task but not women's perception of the gender stereotype of the task.

In the previous section we saw that women's perception of the task's gender stereotype did not differ between both frames. The share of women who considered the task as favoring women was 62% under the *creativity* frame and 58% under the *intelligence* frame. If the gender stereotype was a main driver for competitive performance, then we would *not* expect to see women's performance to be better under the *intelligence* frame. However, this is what we observe. Women performed significantly better in the tournament under the *intelligence* frame than under the *creativity* frame (creativity: 4.75 vs. intelligence: 5.27, $p = 0.047$).

A similar argument holds for the relationship between men's tournament entry decisions and their beliefs about others' perception of the task. We saw that 40% of men under the *creativity* frame but only 29% of men under the *intelligence* frame believed that others consider the task as favoring women. If a man's belief about the task perception of others was a main driver for his tournament entry decision, then we would expect to see more men to enter the tournament under the *intelligence* frame. However, this is not what we observe. Men's likelihood to enter the competition was even slightly higher under the *creativity* frame (creativity: 0.43 vs. intelligence: 0.41, $p = 0.788$).

We could, therefore, not replicate an effect of a task's gender stereotype on men's or women's competitiveness.

2.4 Conclusion

In our experiment we changed the frame of a task to test whether the perceived domain of the task affects man's and women's competitiveness. We find that our participants performed significantly better in the tournament when the task was framed as an *intelligence* task and not as a *creativity* task. To our surprise, the framing effects did not differ between both genders. Men and women performed equally well in the tournament, and decisions to either enter a competition or to select a non-competitive alternative were not affected. Furthermore, the perceived gender stereotype of a task does not seem to affect competitiveness.

These are interesting results for policymakers. If they want to decrease the gender gap in competitiveness, then measures to dismantle gender stereotypes will not be effective. Furthermore, institutions that perform competitive entrance exams do not need to worry about the framing of the tasks. Referring to the exam tasks as "intelligence" tasks or as "creativity" tasks might affect participants' competitive performance but not the gender difference in competitiveness.

Finally, our paper highlights the benefits of experimental designs that change the framing of a task to study the effects of gender stereotypes. Replacing an entire task as a treatment likely affects more than just the perception of a task. If task specific characteristics and not gender stereotypes affect the experimental outcome, researchers might draw incorrect conclusions about the effects of stereotypes. With our experimental design, potentially misleading effects of task specific characteristics are ruled out and more accurate conclusions about the effects of stereotypes can be drawn.

2.A Instructions

Note that the depicted practice exercise below does not show the actual Raven matrix that we used in the experiment.

Welcome!

In the experiment today you will be asked to complete five different tasks. None of these will take more than 4 minutes. At the end of the experiment you will receive $4 for having completed the five tasks. In addition we will randomly select one of the tasks and pay you based on your performance in that task. Once you have completed the five tasks, we will determine which task counts for payment by randomly drawing a number between 1 and 5. Before each task we will describe in detail how your payment is determined within that task.

Your total earnings from the experiment are the sum of: (1) your payment from the randomly selected task, (2) the payment that you earn from your answers in the questionnaire, and (3) your $4 fixed payment for completing all the tasks.

< Next >

Consent form and data protection information

The University of Vienna, Universitaetsring 1, 1010 Vienna, Austria is conducting a scientific study using Prolific today.

Your responses will be recorded on our server. The data generated in the study will be separated from the data in the Prolific system after the payment has been completed. The final dataset will not allow any inference to be drawn about the individual participants or the responses of those participants that can be linked to their identity in any way. Correspondingly, the analysis and presentation of all results of this experiment will be anonymized. The anonymous research data will be archived and will possibly be made available to other scientists for further use.

We want to ask for your consent to participate in the experiment and to the implied processing of data. You can revoke your consent at any time, for example, by sending an e-mail to dominik.stelzeneder@univie.ac.at. Your consent is at the same time the legal basis for the processing of data.

For any general information about the processing of data and data protection, especially about your rights to disclosure, correction, deletion, limitations of processing of data, objection, data transmission and complaints to a supervisory body, please contact the data protection officer of University of Vienna: verarbeitungsverzeichnis@univie.ac.at. You also have the right of complaint to the Austrian Data Protection Agency (dsb@dsb.gv.at).

Participation in today's study is entirely voluntary. You have the possibility to exit the study at any time.

You cannot participate in this experiment without consenting.

I have read the data protection information and I consent to participate in the experiment and to the described data processing procedures:

- ○ Yes
- ○ No

Prolific ID

Please enter your Prolific ID:

< Next >

Creativity Task - Exercise

In each of the first three tasks of the experiment, you will be asked to work on tasks in which you solve problems to earn money. We refer to these tasks as *Creativity Tasks*. We will show you matrices which contain either six or nine separate figures that are related to one another in a particular way.

In each matrix, one of the nine figures is missing (the one in the bottom right). Your task is to detect the pattern in the matrix and to select the correct option from the options numbered 1 to 8 in order to complete the matrix.

Each matrix has exactly one correct solution. You will be given 2 minutes in each task to solve a series of these problems. When you enter an answer, the computer will tell you on the next screen whether your answer was correct or not.

Here is a practice exercise:

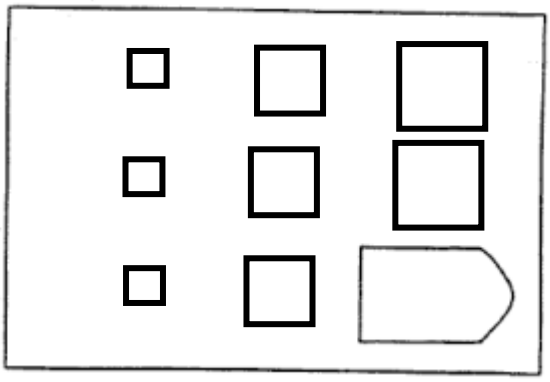
Intelligence Task - Exercise

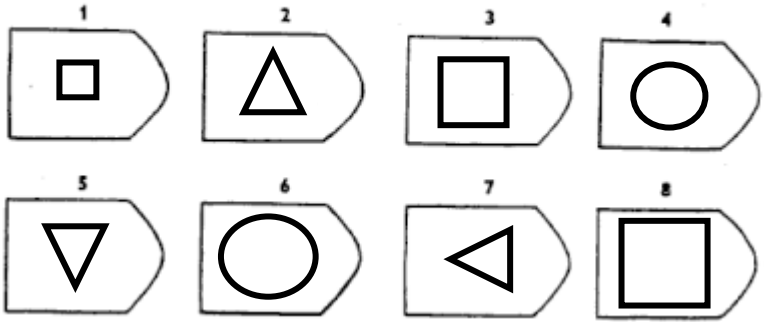
In each of the first three tasks of the experiment, you will be asked to work on tasks in which you solve problems to earn money. We refer to these tasks as *Intelligence Tasks*. We will show you matrices which contain either six or nine separate figures that are related to one another in a particular way.

In each matrix, one of the nine figures is missing (the one in the bottom right). Your task is to detect the pattern in the matrix and to select the correct option from the options numbered 1 to 8 in order to complete the matrix.

Each matrix has exactly one correct solution. You will be given 2 minutes in each task to solve a series of these problems. When you enter an answer, the computer will tell you on the next screen whether your answer was correct or not.

Here is a practice exercise:





Which of the eight options do you choose?

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

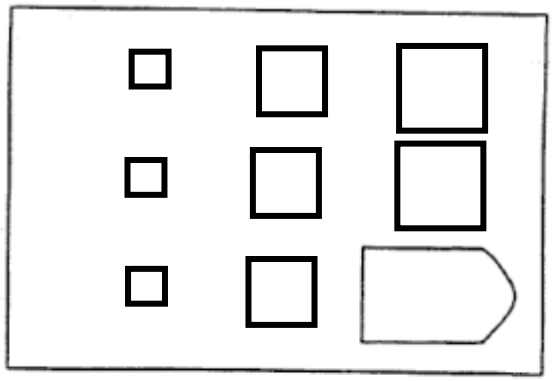
Please select one option and click on the "Next" button. You will see the correct solution on the next page.

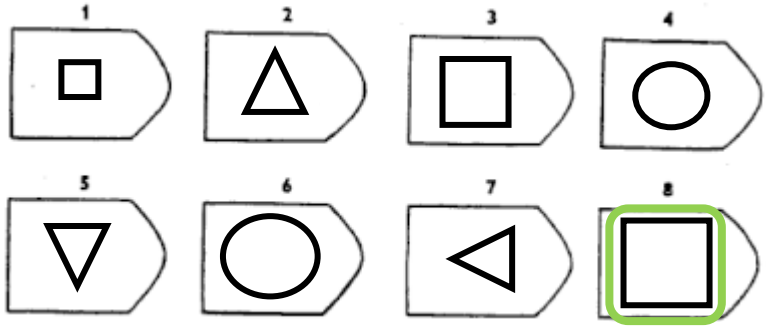
< Next >

Creativity Task - Exercise

Your answer was correct.
Total number of correctly solved matrices in this task: 1

The correct answer in this exercise task is 8.





< Next >

https://de.123rf.com/photo_9934554_colorful-tree.html

Creativity Task 1 - Piece Rate

In *Creativity Task 1* you will be given 2 minutes to solve a series of matrices. If *Creativity Task 1* is the one randomly selected for payment, then you get $0.25 per problem you solve correctly in the 2 minutes. Your payment does not decrease if you provide an incorrect answer to a problem. We refer to this payment as the *piece rate* payment.

Creativity Task 1 will start on the next page.

< Next >

iStock.com/pavlen

Intelligence Task 1 - Piece Rate

In *Intelligence Task 1* you will be given 2 minutes to solve a series of matrices. If *Intelligence Task 1* is the one randomly selected for payment, then you get $0.25 per problem you solve correctly in the 2 minutes. Your payment does not decrease if you provide an incorrect answer to a problem. We refer to this payment as the *piece rate* payment.

Intelligence Task 1 will start on the next page.

< Next >

Creativity Task

https://de.123rf.com/photo_9934554_colorful-tree.html

Remaining time for this task: **1:57**

< Raven Matrix 1 shown here >

< Next >

Intelligence Task

iStock.com/pavlen

Remaining time for this task: **1:57**

< Raven Matrix 1 shown here >

< Next >

Creativity Task

https://de.123rf.com/photo_9934554_colorful-tree.html

Remaining time for this task: **1:39**

Your answer was correct.
Total number of correctly solved matrices in this task: 1

< Raven Matrix 2 shown here >

< Next >

Intelligence Task

iStock.com/pavlen

Remaining time for this task: **1:39**

Your answer was correct.
Total number of correctly solved matrices in this task: 1

< Raven Matrix 2 shown here >

< Next >

https://de.123rf.com/photo_9934554_colorful-tree.html

Creativity Task 2 - Tournament

As in *Creativity Task 1* you will be given 2 minutes to solve a series of matrices. However, for this task your payment depends on your performance relative to that of a group of other participants. Each group will consist of four randomly chosen participants in this experiment. If *Creativity Task 2* is the one randomly selected for payment, then your earnings will depend on the number of problems you solve compared to the three other people in your group. The individual who correctly solves the most problems will receive $1 per correct problem, while the other participants will receive no payment from this task. We refer to this as the *tournament payment*. You will not be informed of how you did in the tournament until all five tasks have been completed. If there is a tie for who solved the most problems, the winner will be randomly determined.

Creativity Task 2 will start on the next page.

< Next >

Creativity Task

https://de.123rf.com/photo_9934554_colorful-tree.html

Remaining time for this task: **1:56**

< Raven Matrix shown here >

iStock.com/pavlen

Intelligence Task 2 - Tournament

As in *Intelligence Task 1* you will be given 2 minutes to solve a series of matrices. However, for this task your payment depends on your performance relative to that of a group of other participants. Each group will consist of four randomly chosen participants in this experiment. If *Intelligence Task 2* is the one randomly selected for payment, then your earnings will depend on the number of problems you solve compared to the three other people in your group. The individual who correctly solves the most problems will receive $1 per correct problem, while the other participants will receive no payment from this task. We refer to this as the *tournament payment*. You will not be informed of how you did in the tournament until all five tasks have been completed. If there is a tie for who solved the most problems, the winner will be randomly determined.

Intelligence Task 2 will start on the next page.

< Next >

Intelligence Task

iStock.com/pavlen

Remaining time for this task: **1:56**

< Raven Matrix shown here >

https://de.123rf.com/photo_9934554_colorful-tree.html

Creativity Task 3 - Payment Scheme Decision

As in the previous two tasks, in *Creativity Task 3* you will be given 2 minutes to solve a series of matrices. However, before doing so, you will now get to choose which of the previous two payment schemes you would like to apply for *Creativity Task 3*. In other words, you can choose whether you would like to be paid in the same way as in *Creativity Task 1* (piece rate payment) or in the same way as in *Creativity Task 2* (tournament payment).

If *Creativity Task 3* is the one randomly selected for payment, then your earnings for this task are determined as follows. If you choose the piece rate payment, you will receive $0.25 per problem you solve correctly.

If you choose the tournament payment, your performance will be compared to the performance of the same three participants who were in your group in *Creativity Task 2*. To decide whether you win if you select the tournament payment in *Creativity Task 3*, your performance score will be compared to the scores of your group members from *Creativity Task 2* (i.e., their past scores). If you correctly solve more problems now than all three of your group members did in *Creativity Task 2*, then you will receive four times the payment from the piece rate, which is $1 per correct problem. You will receive no earnings for this task if you choose the tournament payment and do not solve more problems correctly, than the others in your group did in *Creativity Task 2*. If there is a tie for who solved the most problems, the winner will be randomly determined.

iStock.com/pavlen

Intelligence Task 3 - Payment Scheme Decision

As in the previous two tasks, in *Intelligence Task 3* you will be given 2 minutes to solve a series of matrices. However, before doing so, you will now get to choose which of the previous two payment schemes you would like to apply for *Intelligence Task 3*. In other words, you can choose whether you would like to be paid in the same way as in *Intelligence Task 1* (piece rate payment) or in the same way as in *Intelligence Task 2* (tournament payment).

If *Intelligence Task 3* is the one randomly selected for payment, then your earnings for this task are determined as follows. If you choose the piece rate payment, you will receive $0.25 per problem you solve correctly.

If you choose the tournament payment, your performance will be compared to the performance of the same three participants who were in your group in *Intelligence Task 2*. To decide whether you win if you select the tournament payment in *Intelligence Task 3*, your performance score will be compared to the scores of your group members from *Intelligence Task 2* (i.e., their past scores). If you correctly solve more problems now than all three of your group members did in *Intelligence Task 2*, then you will receive four times the payment from the piece rate, which is $1 per correct problem. You will receive no earnings for this task if you choose the tournament payment and do not solve more problems correctly, than the others in your group did in *Intelligence Task 2*. If there is a tie for who solved the most problems, the winner will be randomly determined.

After the experiment is completely over, we will inform you about how you did relative to the *Creativity Task 2* performance of the others in your group, independent of the payment scheme you choose.

Which payment scheme do you want to apply?

- Piece Rate Payment
- Tournament Payment

Please choose your payment scheme and click "Next". Task 3 will start on the next page.

< Next >

After the experiment is completely over, we will inform you about how you did relative to the *Intelligence Task 2* performance of the others in your group, independent of the payment scheme you choose.

Which payment scheme do you want to apply?

- Piece Rate Payment
- Tournament Payment

Please choose your payment scheme and click "Next". Task 3 will start on the next page.

< Next >

Task 4 – Assessment of Own Past Performance

We would like to ask you how well you think you performed in *Creativity (Intelligence) Task 2* compared to other participants in this experiment (recall that in *Creativity (Intelligence) Task 2* you are matched with three other randomly selected participants and only the best performing participant in your group will receive a payment).

In *Task 4* you will have to answer the following question: "How likely do you think it is that you won the tournament in *Creativity (Intelligence) Task 2*?"

You have the chance to receive a payment of $4 in this task. Note that the incentives are designed such that you maximize your expected payment if you report your true belief.

You can read the details of how your answer to this question will determine your chances of winning $4 by clicking here.

> *After you have entered a number, x, the computer will randomly choose another number between 0 and 100, called y. If this randomly chosen number, y, is larger than the number that you entered, x, then you will receive $4 with probability equal to the value of the randomly chosen number (in percentage points).*
>
> *If the randomly chosen number, y, is smaller than or equal to the number that you entered, x, then your Creativity (Intelligence) Task 2 performance relative to the Creativity (Intelligence) Task 2 performance of the three other participants in your group will determine your payoff in Task 4. If this is the case, then you will receive $4 if you have solved more problems in Creativity (Intelligence) Task 2 than the other three participants of your group and $0 otherwise. In case of a tie the winner will be randomly determined.*

How likely do you think it is that you won the tournament in *Creativity (Intelligence) Task 2*?

Please choose a number between 0 (you are certain that you did not win) and 100 (you are certain that you did win).

The likelihood that I won the tournament in Task 2 is … %

< Next >

Task 5 - Table

In *Task 5* you will make 20 decisions. For each of these decisions, you will be choosing between a payment made with certainty (option A) and a lottery (option B). Option A is identical for every decision problem: you will receive \$0.5 for certain. For option B you will receive either \$2.0 or \$0. The percentage chance with which you will receive \$2.0 will differ for every decision problem.

If *Task 5* is the one randomly selected for payment, then your earnings for this task are determined in the following way:

- First, one of your 20 decisions will be chosen at random.
- In case you chose the certain payment (option A) on the selected decision problem, you will receive \$0.5. In case you chose the lottery (option B), the computer will draw a random number from 0 to 100. If the randomly chosen number is smaller or equal to the probability of receiving \$2.0 in the selected decision problem, you will receive the \$2.0. Otherwise, you will receive \$0. Please have a look at the two examples below.

Example 1:

Decision problem 17 was selected for payment. You chose the lottery (option B).

Row	Option A	Option B
17	\$0.5	20% chance of obtaining \$2.0 80% chance of obtaining \$0

In this example, a random number between 1 and 20 would yield a payment of \$2.0, a higher number (between 21 and 100) would give you a payment of \$0.

Example 2:

Decision problem 12 was selected for payment. You chose the lottery (option B).

Row	Option A	Option B
13	\$0.5	40% chance of obtaining \$2.0 60% chance of obtaining \$0

In this example, a random number between 1 and 40 would yield a payment of \$2.0, a higher number (between 41 and 100) would give you a payment of \$0.

< Next >

Task 5 - Table

Please choose either option A or option B for each row below.

Row	Option A		Option B	
1	○	$0.5	100% chance to obtain $2.0 0% chance to obtain $0	○
2	○	$0.5	95% chance to obtain $2.0 5% chance to obtain $0	○
3	○	$0.5	90% chance to obtain $2.0 10% chance to obtain $0	○
4	○	$0.5	85% chance to obtain $2.0 15% chance to obtain $0	○
5	○	$0.5	80% chance to obtain $2.0 20% chance to obtain $0	○
6	○	$0.5	75% chance to obtain $2.0 25% chance to obtain $0	○
7	○	$0.5	70% chance to obtain $2.0 30% chance to obtain $0	○
8	○	$0.5	65% chance to obtain $2.0 35% chance to obtain $0	○
9	○	$0.5	60% chance to obtain $2.0 40% chance to obtain $0	○
10	○	$0.5	55% chance to obtain $2.0 45% chance to obtain $0	○
11	○	$0.5	50% chance to obtain $2.0 50% chance to obtain $0	○
12	○	$0.5	45% chance to obtain $2.0 55% chance to obtain $0	○

11	○	$0.5	50% chance to obtain $2.0 50% chance to obtain $0	○
12	○	$0.5	45% chance to obtain $2.0 55% chance to obtain $0	○
13	○	$0.5	40% chance to obtain $2.0 60% chance to obtain $0	○
14	○	$0.5	35% chance to obtain $2.0 65% chance to obtain $0	○
15	○	$0.5	30% chance to obtain $2.0 70% chance to obtain $0	○
16	○	$0.5	25% chance to obtain $2.0 75% chance to obtain $0	○
17	○	$0.5	20% chance to obtain $2.0 80% chance to obtain $0	○
18	○	$0.5	15% chance to obtain $2.0 85% chance to obtain $0	○
19	○	$0.5	10% chance to obtain $2.0 90% chance to obtain $0	○
20	○	$0.5	5% chance to obtain $2.0 95% chance to obtain $0	○

Next

Thank You!

Here and on the next pages, we kindly ask you to answer some additional questions. Please select the option *index finger* below. Thanks for reading our instructions!

Based on the text you read above, which finger have you been asked to select?

- ○ little finger
- ○ ring finger
- ○ middle finger
- ○ index finger
- ○ thumb

< Next >

Questionnaire I

Please answer the following questions:

What is your gender?

- ○ Woman
- ○ Man
- ○ Prefer not to say

What is your age?

Please think back to *Creativity (Intelligence) Tasks* 1 to 3 in which you were asked to select the correct missing figure among the proposed options.

Have you ever seen this kind of task before?

Have you ever solved this kind of task before?

< Next >

Questionnaire II

Who do you think performed better in *Creativity (Intelligence) Task 2* (the tournament) on average? Men or women?

You will receive $0.25 if your answer is correct.

- ○ Men performed better in *Creativity (Intelligence) Task 2* on average
- ○ Women performed better in *Creativity (Intelligence) Task 2* on average

How much did you enjoy working on *Creativity (Intelligence) Tasks 1 to 3* on a scale from 0 (not at all) to 10 (very much)?

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

What do most people think, are men or women better at solving tasks like those in *Creativity (Intelligence) Tasks 1 to 3*?

You will receive $0.25 if you choose the same answer as the majority of participants in this experiment.

- ○ Most people think that men are better at solving tasks like those in *Creativity (Intelligence) Tasks 1 to 3*.

- ○ Most people think that women are better at solving tasks like those in *Creativity (Intelligence) Tasks 1 to 3*.

< Next >

Questionnaire III

For each of the following items, please choose either A or B. One of the items will be randomly selected for payment at the end of the experiment.

Item	Option A		Option B	
1	○	10% chance to obtain $0.50 90% chance to obtain $0.40	10% chance to obtain $0.95 90% chance to obtain $0.05	○
2	○	20% chance to obtain $0.50 80% chance to obtain $0.40	20% chance to obtain $0.95 80% chance to obtain $0.05	○
3	○	30% chance to obtain $0.50 70% chance to obtain $0.40	30% chance to obtain $0.95 70% chance to obtain $0.05	○
4	○	40% chance to obtain $0.50 60% chance to obtain $0.40	40% chance to obtain $0.95 60% chance to obtain $0.05	○
5	○	50% chance to obtain $0.50 50% chance to obtain $0.40	50% chance to obtain $0.95 50% chance to obtain $0.05	○
6	○	60% chance to obtain $0.50 40% chance to obtain $0.40	60% chance to obtain $0.95 40% chance to obtain $0.05	○
7	○	70% chance to obtain $0.50 30% chance to obtain $0.40	70% chance to obtain $0.95 30% chance to obtain $0.05	○
8	○	80% chance to obtain $0.50 20% chance to obtain $0.40	80% chance to obtain $0.95 20% chance to obtain $0.05	○
9	○	90% chance to obtain $0.50 10% chance to obtain $0.40	90% chance to obtain $0.95 10% chance to obtain $0.05	○
10	○	100% chance to obtain $0.50 00% chance to obtain $0.40	100% chance to obtain $0.95 00% chance to obtain $0.05	○

Questionnaire IV

Please choose one of the following lotteries. Your chosen lottery will be paid out to you at the end of the experiment.

- 50% chance to obtain $0.70; 50% chance to obtain $0.70
- 50% chance to obtain $0.60; 50% chance to obtain $0.90
- 50% chance to obtain $0.50; 50% chance to obtain $1.10
- 50% chance to obtain $0.40; 50% chance to obtain $1.30
- 50% chance to obtain $0.30; 50% chance to obtain $1.50
- 50% chance to obtain $0.05; 50% chance to obtain $1.75

< Next >

Thanks a lot for participating in this experiment!

On the next page we will inform you about your earnings.

< Next >

Thanks for your participation!

The computer has randomly selected Task 1 to be payment relevant for you. Your payoff is the sum of the following payments.

- You will receive $4 for having completed the five tasks.
- In Task 1 you solved 3 matrices. You receive $0.25 for each correct solution and hence will receive $0.75 in total for your performance in Task 1.
- Your earnings for your answers in questionnaire II depend on the performance and on the decisions of other participants of this experiment. We thus cannot inform you about your earnings in questionnaire II now. We will send you a message with information about your earnings in this questionnaire after the experiment has ended.
- For questionnaire III row 9 was randomly selected and hence your payoff increases by $0.5.
- For your choice in questionnaire IV, you will receive an additional payment of $0.05.

In total you will receive $4 for having completed the five tasks and a bonus of $0.75 for Task 1 plus $0.5 and $0.05 for questionnaires III and IV. In addition, we will inform you soon in a separate message about the bonus that you will receive from questionnaire II.

Chapter 3

Gender Differences in Bayesian Updating: Experimental Evidence[1]

Misinformation in consumer markets - like fake reviews or biased endorsements - poses a challenge for consumers that can be conceptualized as a two-step problem. First, they need to form a belief about the probability that the information is biased. Second, they need to update their beliefs conditional on receiving information. This paper reports on an experiment that disentangles both steps utilizing a novel approach which relies neither on the estimation of individual utility parameters nor on belief elicitation. We find that inability to correctly compute posterior probabilities significantly impairs decisions which leads to a reduction in consumer surplus and that this effect is heterogeneous across gender.

3.1 Introduction

Situations of asymmetric information arise in everyday interactions and are relevant in many market contexts. In consumer markets, for instance, producers or sellers typically hold more information about a product than consumers. The emergence of customer reviews ostensibly reduces this asymmetry through dissemination of information among consumers. However, the presence of fake reviews introduces another dimension of information asymmetry. Not only are consumers uninformed about product quality but, in addition, also about the credibility of the very reviews they consult in order to inform themselves.[2] This results in a two-step problem for consumers. First, they need to form a belief about the trustworthiness of a review. Second, they need to map this belief to-

[1] This chapter is joint work with Boris Knapp.

[2] Mayzlin, Dover, and Chevalier (2014) exploit organizational differences between two platforms to show empirically that reviews differ significantly when faking them is made more difficult. He, Hollenbeck, and Proserpio (2021) directly track down recruitment of fake reviewers via Facebook.

gether with the review's information into a posterior belief about product quality. Similar situations arise in other contexts as well. Consider for example a hiring decision, where the applicant has private information about their ability while the recruiter needs to infer it from a letter of reference. Sometimes these letters are not honestly assessing candidates' abilities but instead trying to portrait them in a favourable light. Thus, on top of the asymmetry in information about the candidate's type, there is additional uncertainty about the trustworthiness of the source of information which is supposed to resolve the former uncertainty. Financial markets exhibit a similar problem of asymmetric information because credit rating agencies (CRAs) have been found to inflate ratings in order to foster business relations (Fulghieri, Strobl, and Xia 2014).

Therefore, while reviews, letters of reference and credit ratings are instruments for resolving information asymmetries, fake reviews, dishonest reference letters, and inflated ratings are detrimental to this objective. Addressing the problem of misinformation directly is a natural approach to the problem. Competition (Board and Lu 2018, Au and Kawai 2020, Au and Kawai 2021), reputation (Sobel 1985, Ottaviani and Sørensen 2006), or regulation (Bourjade and Jullien 2011, Janssen and Roy 2022) are all means to reduce misinformation but cannot alleviate the problem completely. Even though the credit rating market is competitive and CRAs have reputation concerns, inflated ratings can still be observed. Despite Amazon's introduction of a verification system and its enormous effort in fighting fake reviews, a substantial share of reviews continues to be fake.[3]

It is thus necessary to understand and mitigate the challenge that the existence of misinformation imposes on consumers. Since the presence of misinformation makes truthful information less informative, even a fully rational decision maker who perfectly understands the problem suffers from it. However, cognitive biases and limitations may exacerbate this problem. First, decision makers' beliefs about the trustworthiness of information might be biased or noisy. If consumers believe the share of fake reviews to be higher or lower than it actually is, this might affect the inference they make negatively. Second, misinformation complicates the inference process as it becomes a (more) complex Bayesian updating problem. Mistakes in Bayesian updating might thus have additional negative effects and consumer welfare could decrease even more as a result.

In this paper, we study these negative effects and how to mitigate them. We do this by mimicking informational policies that can be implemented relatively easily and already have been to some extent, for example in the context of online reviews.[4] In particular,

[3]According to the monitoring service *Fakespot*, 42% of 720 million assessed Amazon reviews in 2020 were allegedly fake. See `https://www.chicagotribune.com/business/ct-biz-amazon-fake-reviews-unreliable-20201020-lfbjdq25azfdpa3iz6hn6zvtwq-story.html`, retrieved on June 9, 2022.

[4]Fakespot.com provides an estimate of the share of fake reviews together with a corrected rating for a

we focus on two interventions, each targeting one of the two potential sources of errors described above, namely misspecified beliefs about the trustworthiness of information and mistakes in Bayesian updating. To this end, we conduct an experiment where subjects need to condition their decisions based on potentially misleading information, reflecting a situation of information asymmetry as described above and common in many market contexts.

Our experiment is based on a simple signaling game (Crawford and Sobel 1982), where a sender sends an envelope that contains either \$5 or \$0 with equal probability to a receiver. Only the sender observes the content of the envelope and, before sending it to the receiver, he labels it with one of two messages: "This envelope contains \$5" or "This envelope is empty". Based on this label and knowing that it is not necessarily truthful, the receiver then decides whether to open the envelope or to take a fixed payment instead (her outside option). This decision affects the payoffs of both players. The sender earns a positive payoff only if the receiver opens the envelope and nothing otherwise. The receiver gets her outside option if she decides not to open the envelope and the content of the envelope (either \$5 or \$0) otherwise.

Two features of our experiment allow us to assess the extent to which the receiver deviates from her individually optimal decision. First, we use the Becker, Degroot, and Marschak (1964) mechanism to elicit subjects' valuation of the envelope.[5] Second, in an additional task we elicit the receiver's *actual valuation* of the envelope, i.e. her valuation if she would have (i) known the exact likelihood that the label was wrong and (ii) been perfectly able to apply Bayes' rule. The difference between a subject's valuation in the first and second task defines the *decision error*.

In our experiment, we study how two informational treatments affect receivers' decision errors and implied gains in consumer welfare. In the first treatment $T1$ (*correct beta*), the receivers are informed about the true probability that a label is misleading, which mitigates potential negative effects due to misspecified beliefs. In our second treatment $T2$ (*slider*), we provide receivers with a "slider" that maps any prior belief into a posterior probability and thus reduces possible adverse consequences of incorrect Bayesian updating. Treatment one thus exogenously induces the correct belief and treatment two enables participants to correctly perform Bayesian updating. In a third treatment $T3$ (*full info*), we study the joint effect of these policies by providing them with all information of the first and second treatment.

Based on the previous experimental literature, a clear prediction for the effects of our treatments cannot be made. The effects of the first treatment will depend on the

given product.

[5]Before learning the value of her outside option, the receiver had to indicate the smallest outside option that she would be willing to accept for not opening the envelope.

accuracy of the receivers' beliefs about the share of senders that send the honest message. We expect treatment T1 to have a larger effect if the belief accuracy is low and a smaller or even no effect if the beliefs are accurate. The literature finds mixed results for the accuracy of subjects' beliefs. While Aoyagi, Fréchette, and Yuksel (2022) assess the beliefs of subjects in their experiment as accurate, Costa-Gomes and Weizsäcker (2008) and Rey-Biel (2009) find substantial inaccuracies of stated beliefs. The different observations in those papers might, however, be the result of different games that were used in those papers. The former study used a repeated prisoners' dilemma game, whereas in the other two papers behavior in various one-shot 3 x 3 games was analyzed. Since our experiment differs substantially from those three experiments in the literature, no clear prediction for the effects of treatment $T1$ (*correct beta*) can be made. Concerning the effects of the second treatment, the literature offers some guidance as to what can be expected. Benjamin (2019) reports a number of studies that document biases in Bayesian updating (e.g. Phillips and Edwards 1966 and Edwards et al. 1968 for two early papers). The provision of a slider that maps any prior belief into a posterior probability helps to prevent these biases. We thus expect treatment $T2$ (*slider*) to decrease subjects' decision errors and thus have a significant negative effect on the error size. Regarding gender differences, Jendraszek (2008) studies probabilistic reasoning among future mathematics teachers and finds that men outperform women in tasks that involve conditional probability problems. We thus hypothesize that women benefit more from treatment $T2$ than men.

There is a large and growing literature that studies fallacies in Bayesian updating and the role of beliefs in economic decision making (Benjamin 2019 and Nyarko and Schotter 2002). Our contribution to this literature is fourfold. First, despite extensive research on the role of beliefs and Bayesian updating in economic decision making, we are the first to study the effect of both roles jointly. Our research thus informs policy makers and organizations that aim to improve customers' purchasing decisions. Consider, for instance, the website fakespot.com that checks the reviews for a given listing in an online store like Amazon and estimates the share of fake reviews. Additionally, it computes a corrected rating taking the estimated share of fake reviews into account. It is important to notice that these two pieces of information differ significantly in the cost of provision. Estimating the share of fake reviews typically involves complex language analysis and machine learning techniques and thus requires substantial resources. In contrast, computing a new rating based on an estimated share of fake reviews is a computation that can be done relatively easily and inexpensively. Understanding the extent to which both these measures help consumers is therefore valuable information for implementing such policies.

Second, we contribute to the literature on gender differences in probabilistic reason-

ing. Many markets are segmented according to gender or exhibit a high concentration of one gender. In such markets, taking into account gender differences is crucial for efficient policy design. And even in markets that are neither segmented nor dominated by one of the genders, understanding gender differences is important for assessing how a certain policy might affect men differently than women. Even though probabilistic reasoning has been studied extensively in the aggregate and gender differences across many other dimensions have gathered attention, gender differences in probabilistic reasoning remain relatively understudied. With the recent surge of information economics and the continuing focus on gender differences not only in economics but also other social sciences, this comes as somewhat of a surprise. To our knowledge, the only study specifically exploring gender differences in probabilistic reasoning is Jendraszek (2008), who studied to what extent future teachers of mathematics were subject to a range of probabilistic misconceptions. She finds that men outperformed women in tasks involving conditional probability problems. This suggests that Bayesian updating might impair women's decision making to a greater extent. In our paper we study whether male or female subjects benefit more from a tool that facilitates Bayesian updating.

Third, we add to a growing literature that quantifies individual decision errors and consequent losses in consumer surplus (Carpenter et al. 2021, Harrison, Morsink, and Schneider 2020, and Harrison et al. 2020). In this literature, individual utility parameters are structurally estimated based on individuals' pretreatment choices. The actual decisions of interest are then compared to the benchmark decisions that maximize the individuals' specific utility functions. In our paper, we *directly* elicit the individually optimal decision of each participant and we compare this decision to the decision of interest. Our conclusions about individuals' decision errors and foregone consumer surplus thus do not depend on a presumed form of utility function. Importantly, this approach allows to estimate individual decision errors and not only systematic biases on a population level.

Our fourth contribution is to the literature that studies the role of beliefs in games, which can be broadly categorised into two groups. One group of papers estimates beliefs from observed choices (Schneider 2019, Aguirregabiria 2021). The other group of papers elicits beliefs and then relates them to subsequent decisions (Nyarko and Schotter 2002, Costa-Gomes and Weizsäcker 2008). In our paper, we use a treatment that directly induces the correct belief instead. We then compare subsequent decisions to the decisions of a control group of participants that was not treated. Estimating or eliciting subjects' beliefs is thus not required. Consequently, a clearer causal link between beliefs and subsequent decisions can be established because, for example, specific assumptions about subjects' utility functions are not required.

The results of our experiment are surprising in light of the related literature and have

important implications for designing informational policies. While we do not find significant treatment effects in the aggregate sample, analyzing men and women separately reveals unexpected differences. Looking at the male subsample, we find that the *full info* treatment $T3$ improved decisions significantly and substantially, reducing the decision error by 34%. This reduction seems to be largely driven by the improvement in Bayesian updating, given that the *slider* treatment $T2$ had an effect of similar size. Treatment $T1$ (*correct beta*) had no significant effect, however, due to the observed carry-over effect of this treatment, we are cautious with interpreting this result. Focusing on the female subsample, we find no significant effects of any of our treatments. Interestingly, the literature about gender differences in Bayesian updating would suggest that the effect of $T2$ should be larger for women than for men. Finally, the effects on expected consumer surplus mirror the ones on decision error. This is to be expected given that the expected consumer surplus is a function of the decision error.

The remainder of the paper is structured as follows. Section 3.2 explains our experimental design. In section 3.3 we introduce the concept of decision error and show how it relates to consumer surplus. In section 3.4, we present the results of the experiment. After first providing some summary statistics, we look into issues that could potentially confound our findings, and finally, we present the treatment effects. Section 3.5 concludes.

3.2 Experimental Design

Our experiment was based on the canonical sender-receiver game. Each sender had private information about the content of an envelope, either \$0 or \$5, and had to send one of two messages to the receiver of the envelope:

A: "This envelope contains \$5."

B: "This envelope is empty."

Subsequently, the receiver decided whether to open it, receiving its content, or sell it for an amount that had been randomly drawn from $[\$0.00, \$5.00]$. The sender received \$3 if the envelope was opened and \$0 otherwise. Both, the senders and the receivers were informed about this structure of the game.

We recruited a total of 200 senders and 400 receivers via *Prolific*. The sample consisted of US citizens and was representative with respect to age, ethnicity and sex. The experiment was programmed in *oTree* (Chen, Schonger, and Wickens 2016) and took place in January 2022.

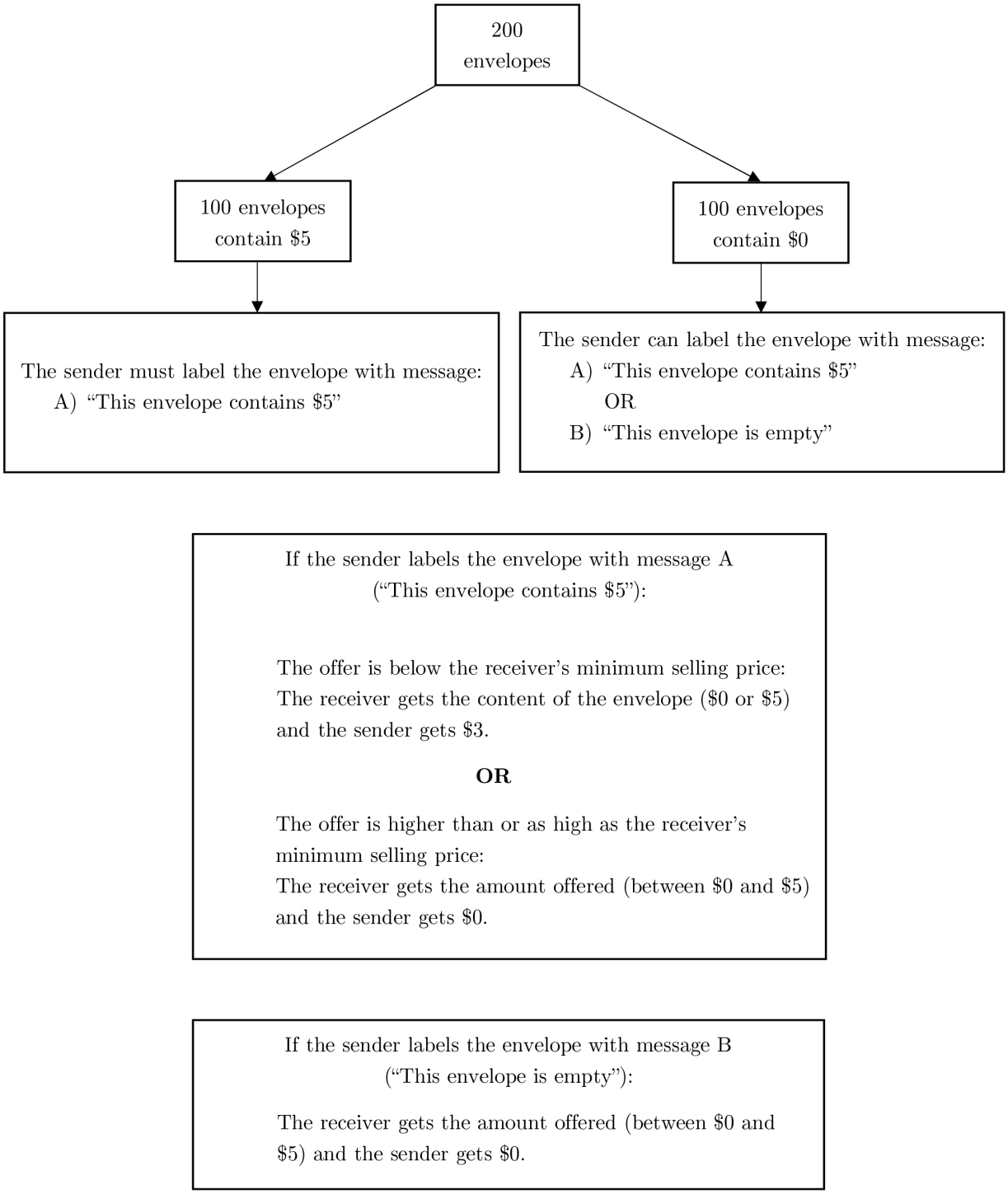


Figure 3.1: Graphical illustration of the sender-receiver game.

3.2.1 Senders

We distributed 200 envelopes - 100 empty, 100 containing \$5 - randomly among 200 senders. Before observing their envelope's content, the senders had to choose between messages A and B for the case that it was empty. In the case that the envelope contained \$5, message A was sent automatically. After they had made a choice, they were informed about their envelope's content and had to fill out a short demographic questionnaire. The envelope, labeled with one of the two messages, was then sent to a randomly selected receiver to be either opened or sold unopened. If the envelope was opened, the sender earned a bonus of \$3 on top of a \$2 participation fee. Otherwise, they received the participation fee but no bonus.

3.2.2 Receivers

The receivers' side of the experiment consisted of a total of six tasks. At the end of the experiment, one of the tasks $1-4$ was randomly selected for payment. On top of that, Tasks 5 and 6 were both paid out.

After having read the instructions and having seen a graphical illustration of the experiment, participants had to answer six comprehension questions about the experiment. The questions helped to convey the instructions because immediately after submitting their answers, the correct answers to all six questions along with short explanations were provided.

In Task 1, each receiver was randomly assigned one of 200 envelopes. They knew that 100 of them contained \$5, while the remaining ones were empty, and that a sender labeled each of the envelopes according to the rules explained above. Thus, the envelope was labeled either with message A ("This envelope contains \$5.") or with message B ("This envelope is empty."). Before finding out which of the two messages was written on their envelope, they had to name a minimum selling price from $[\$0.00, \$5.00]$ for the case that it would be message A.[6] Only then the participants were informed about the label of their envelope. If the envelope was indeed labeled with message A, the computer generated a random amount from $[\$0.00, \$5.00]$ and offered it in exchange for the envelope. If that offer was at least as high as the stated minimum selling price, the envelope was sold and the receiver was awarded with the offered amount. The receiver did not learn about the content of the envelope in this case. Otherwise, if the offer was below the minimum selling price, the envelope was not sold but opened instead. In this case the receiver got its content, either \$0 or \$5. If, however, the message was B, then the envelope was sold

[6] They did not have to specify a minimum selling price for the case that it would be message B. A letter labeled with message B was sold automatically, which is equivalent to a minimum selling price of 0.

automatically and the receiver was awarded with the offered amount. We thus used the Becker-DeGroot-Marschak (BDM, 1964) mechanism to elicit subjects' valuations for an envelope labeled with message A.

In Task 2, each receiver was given a lottery ticket with a winning probability that was based on β, the share of senders who lied. It was calculated to be equal to the posterior probability that an envelope showing message A contained \$5. The outcome of the lottery had two components, a payoff to the receiver and a payoff of \$3 to one sender. If the ticket won, the receiver got \$5 and the sender who got the payoff of \$3 was randomly chosen among those whose envelopes contained \$5. If it lost, the receiver got \$0 and the sender was randomly selected among those whose envelopes contained \$0 and who chose to label them with message A. With the same probability that the envelope in Task 1 showed message B, the lottery ticket was sold automatically. Thus, the lottery was payoff equivalent to Task 1. As in Task 1, we elicited subjects' minimum selling price for the lottery ticket via the BDM method. We did not inform participants about the outcome of a lottery if the ticket was sold.

Task 3 consisted of a lottery similar to the one in Task 2, with the only difference that no sender received any payoff. This lottery was thus equivalent to the lottery in Task 2, except for the social externality. As in Tasks 1 and 2, no feedback was given before the end of the experiment.

Task 4 was a cognitive ability test consisting of two parts. The first part was the *Cognitive Reflection Test* introduced by Frederick (2005). It is comprised of three questions, that have intuitive answers, which, however, are incorrect. Only deeper reflection reveals the correct answer. For each correct answer, the participants earned 40 cents. The second part consisted of a series of *Raven's Progressive Matrices* (Raven, Raven, and Court 1998). The matrices contained either six or nine separate figures that were related to one another in a particular way. In each matrix, one of the nine figures was missing. The participants had to select the correct figure that completed the matrix. Participants had three minutes to complete as many of the 34 matrices as possible and earned 30 cents for every correct solution. Before the timer started, we explained the task with an example.

In Task 5, we used the frequency method (Schlag and Tremewan 2021) to elicit subjects' beliefs about the share of senders who lied. They received a payoff of \$1 if their answer deviated by less than three percentage points from the true β and nothing otherwise.

Task 6 consisted of a series of 20 decision problems. For each of these problems the participants had to choose between a risky lottery and a safe outcome. The lottery was the same in each decision problem, yielding \$1 with a probability of 75% and \$0 with

the remaining probability. The safe outcome was a payoff of 5 cents in the first decision problem and increased by 5 cents for every subsequent problem. Hence, the twentieth problem was a choice between the lottery and a safe payoff of \$1. Of these 20 decisions, one was randomly chosen for payment.

The BDM mechanism used in Tasks $1-3$ is only incentive compatible under the assumption of monotonicity (Azrieli, Chambers, and Healy 2018). With Task 6 we can test for strict monotonicity and disregard participants whose preferences do not satisfy this condition from our analyses. To do so, we count the number of decision problems at which a subject switches from choosing the risky lottery to choosing the safe payoff and vice versa. A participant whose utility preferences satisfy the axiom of strict monotonic preferences will either have one or no such switching point. The reason is that after switching from the lottery to the safe payoff once, all subsequent safe payments are larger and thus necessarily preferred given that the subjects has strict monotonic preferences.

Finally, participants had to fill out a short demographic questionnaire before we informed them about all outcomes and their payoffs in each of the chosen tasks. In addition to what they earned in these tasks, they were paid a participation fee of \$4.[7]

3.2.3 Treatments

Our experiment featured a 2×2 between subject design, varying receivers' Task 1. The motivating idea was that for participants to choose optimally in Task 1, two steps were required. First, they needed to have an accurate belief about β, the share of senders who lied. Second, based on their estimate of β, they needed to calculate the posterior probability that an envelope showing message A indeed contained \$5. Our design varied the information provided in Task 1. The baseline treatment $T0$ was exactly as explained above, receivers had no information besides the prior probability that any given envelope contained \$5.

In treatment $T1$ we informed receivers about the share of senders who lied. To this end, we exploited the sequential nature of our experiment, running the sender part first, thus obtaining the realized β before running the receiver part.

In treatment $T2$ we provided participants with a device for the specific purpose of calculating the posterior probability that their envelope was full, respectively empty, given some belief about β. The device in treatment $T2$ took the form of a slider, which they could move to positions from 0% to 100% representing beliefs about β. The formulas for Bayesian updating that the slider applied were as follows.

$$\pi(\$5|A) = \frac{1}{1+\beta} \qquad (3.1) \qquad\qquad \pi(\$0|A) = \frac{\beta}{1+\beta} \qquad (3.2)$$

[7] They received \$2 more than senders because the experiment took longer for receivers.

Task 1

By clicking on the ‚+' symbols you can access the instructions again. Please take your time and consider your decision carefully.

+ Information on Task 1 (senders)

+ Information on Task 1 (receivers)

+ Graphical Illustration

Additional Informationen:

57% of all senders whose envelopes were empty labelled them with message A ("This envelope contains $5")

Below you can use a calculator that was designed specifically for this task. You can submit an estimate of the share of *senders* who labelled their empty envelopes with message A ("This envelope contains $5"). Based on your estimate, the calculator computes the probability that *your* envelopes contains either $0 or $5. You can submit multiple estimates and use the calculator as often as you wish.

Input:

If 57 % of those *senders* whose envelopes were empty, labelled them with message A ("This envelope contains $5") and you receive an envelope that is labelled with message A, then:

Result:

- Your envelope contains $5 with a probability of 64 %.
- Your envelope contains $0 with a probability of 36 %.

Please answer the following question:

If your envelope is labeled with the message A ("This envelope contains $5"), for how much money would you be willing to sell it to us?

Your answer: $1.8

If we offer you between $0 and $1.79 you will open the envelope.

If we offer you $1.8 or more you will sell the envelope.

Next

Figure 3.2: Screenshot of Task 1 with information on both the share of senders who lied and the slider.

Table 3.1: Treatments

		correct β	
		no	yes
slider	no	*baseline*	*correct beta*
	yes	*slider*	*full info*

Notes. Table shows experimental design.

The slider applied Equation (3.1) to calculate the posterior probability that the envelope contains \$5, given message A and the participant's belief about β. With Equation (3.2), the slider calculated the posterior probability that the envelope was empty, given the same conditions. Below the slider, participants were informed about the outcomes of the calculations. Hence, they could see the probabilities that the envelope was filled or empty, respectively, conditional on the belief at which they positioned the slider.

Treatment $T3$ combined both pieces of information from the previous treatments. Participants were informed about the correct β *and* they could use a slider mapping prior beliefs into posteriors (see Figure 3.2). Thus, $T3$ eliminated the strategic uncertainty and computational complexity of Task 1 and rendered it payoff-equivalent to the lottery in Task 2. Table 3.1 summarizes the treatment design.

3.3 Decision Error and Welfare

Our main research question is whether and how our treatments affect men's and women's decision error in Task 1. Judging the error of a decision requires comparing the actual decision of a participant with the "optimal" decision that he or she "should have made".

Carpenter et al. (2021), Harrison, Morsink, and Schneider (2020), and Harrison et al. (2020) are the first that quantify decision errors and the consequent losses in expected consumer surplus. In their experiments, they use participants' pretreatment choices to structurally estimate the parameters of a presumed utility function for each individual. The actual posttreatment decisions of interest are then compared to the benchmark decisions that maximize the individuals' specific utility functions.

In our paper we use a slightly modified approach that does not require structural estimations or assumptions about the specific form of a utility function but instead provides the individually optimal decision directly. In our study, a participant's decision in Task 2 defines her individually optimal decision for Task 1. Recall that Task 2 was identical to Task 1 in terms of outcome. The only difference was that in Task 2 all relevant information for the decision problem was available, which was not necessarily the case in

Task 1. Under the assumption of consistency, the decision in Task 2 informs us about the actual valuation of the envelope in Task 1. This is also what a participant *should have chosen* as her minimum selling price in Task 1 if the additional information would have been available to her.

The validity of this experimental approach depends on the assumption that effects of our treatments in Task 1 do not carry over to decisions in Task 2. Such carry-over effects are a well-known issue in experimental within-subject designs (Greenwald 1992, Charness, Gneezy, and Kuhn 2012) and we will test for them in section 3.4.2.

The main advantage of our approach compared to other approaches in the literature is that our results are independent of the specific utility functions of our participants. We can assess the size of the decision error by directly comparing the decisions made in Tasks 1 and 2. If both decisions coincide, then the decision in Task 1 was individually optimal and the error was *zero*. If, however, the decisions do not coincide, then the decision in Task 1 contained an error equal to the extent of the deviation. Given the discussion above, we believe that assuming that the decision in Task 1 but not the decision in Task 2 might contain a decision error is well founded.

The economic consequences of such deviations can be illustrated by a simple welfare analysis. We will see that the size of a decision error has direct implications for a participant's expected consumer surplus (CS).

The calculation of a participant's expected CS is straightforward. Let us first assume that the decision error is 0. This means that a subject's valuation of the envelope, x, and her stated minimum selling price (MSP), y, are both equal, i.e. $y = x$. In this case, the expected CS can be calculated as follows: She obtains a surplus of x whenever the randomly drawn offer in exchange for the envelope is below x and the money offered otherwise. Recall that the offer is a randomly drawn amount from the interval $[0, 5]$. Let us denote the offer by z such that $F(z) = \frac{z}{5}$. Then,

$$\begin{aligned} CS_{no\_error} &= \int_0^x x dz + \int_x^5 z dF(z) = F(x)x + \int_x^5 \frac{z}{5} dz = \frac{x^2}{5} + \frac{z^2}{10}\Big|_x^5 \\ &= \frac{2x^2}{10} + \frac{5^2}{10} - \frac{x^2}{10} = \frac{x^2 + 5^2}{10} \end{aligned}$$

Given that a subject's valuation of the envelope lies in $[0, 5]$, CS_{no_error} takes its minimum at $x = 0$ where it is 2.5 and its maximum at $x = 5$ where it is 5. Figure 3.3 illustrates the CS in the case where MSP and valuation coincide, i.e. without decision error.

Let us now turn to the case in which subjects make decision errors by submitting MSPs that are not equal to their valuations of the envelope. As before, x denotes a subject's

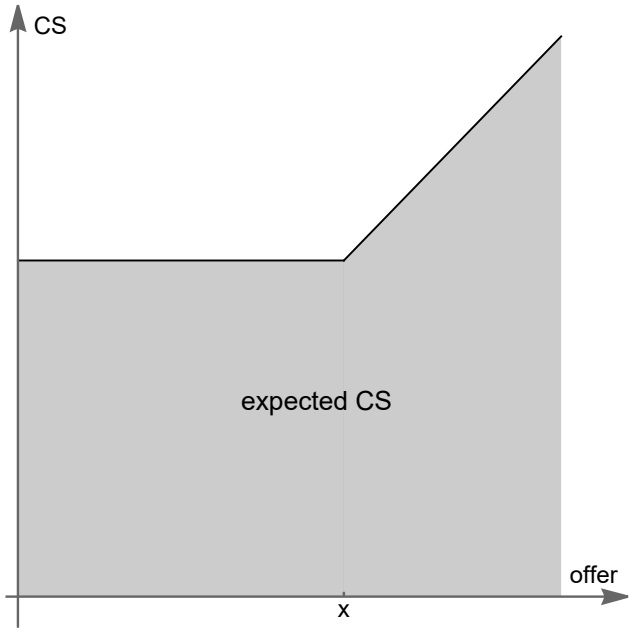


Figure 3.3: Expected Consumer Surplus in the case where the MSP is equal to the valuation.

valuation for the envelope and y their MSP, only now $y \neq x$. Then, the expected CS in Task 1 can be computed as:

$$\begin{aligned} CS_{with\_error} &= \int_0^y x dz + \int_y^5 z dF(z) = F(y)x + \int_y^5 \frac{z}{5} dz = \frac{xy}{5} + \frac{z^2}{10}\Big|_y^5 \\ &= \frac{2xy}{10} + \frac{5^2}{10} - \frac{y^2}{10} = \frac{x^2 - (x-y)^2}{10} + \frac{5^2}{10} = CS_{no\_error} - \frac{(x-y)^2}{10} \end{aligned}$$

And the foregone consumer surplus is simply:

$$CS_{foregone} = \frac{(x-y)^2}{10} \tag{3.3}$$

Thus, the loss in consumer surplus is proportional to the decision error squared. Graphically, this can be illustrated as in Figure 3.4. The left panel depicts the case where the MSP is below the actual valuation of the envelope. In this case, whenever the offer is between the MSP and the valuation, the participant receives y although she would have valued opening the letter with $x > y$. Mathematically, the loss in expected CS is the integral of the difference over all such offers, weighted by their probability densities. This corresponds to the the shaded area. The case where the MSP is above the valuation is depicted in the right panel of Figure 3.4. In both cases it is thus the difference between valuation and MSP, i.e. the *decision error*, that determines the loss in expected CS.

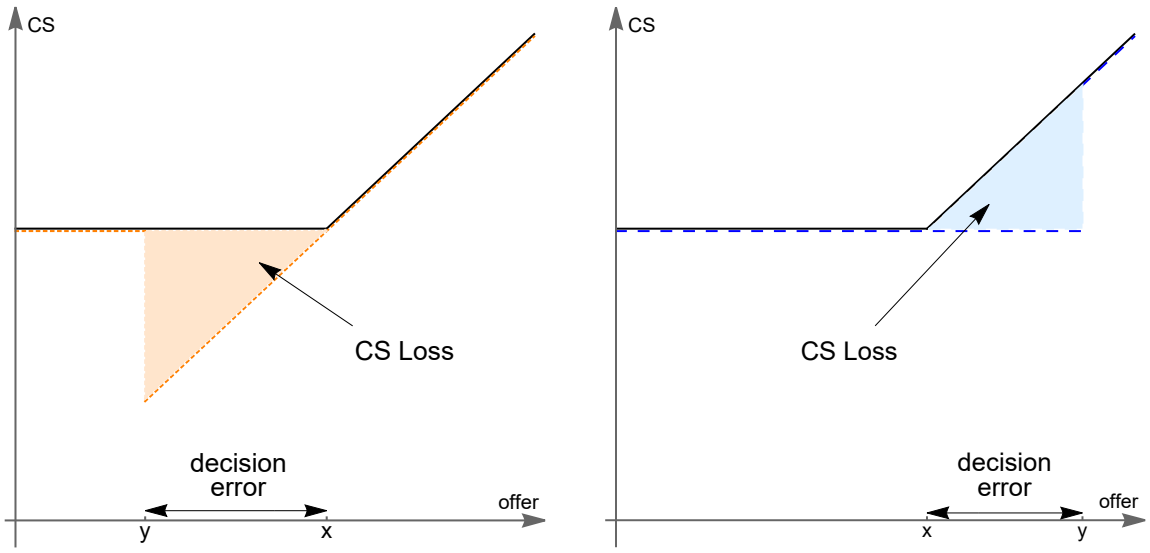


Figure 3.4: The shaded areas illustrate the loss in CS due to decision error. The left(right) panel depicts the case where the MSP is lower(higher) than the valuation.

3.4 Results

We will present the results of this study in the following sequence. In section 3.4.1 we summarize the decisions of the senders and present the implications for the receivers in the main part of the experiment. Section 3.4.2 is concerned with issues that could arise if subjects violate the axiom of monotonic preferences or if treatment effects in Task 1 carry over to Task 2. Finally, section 3.4.3 presents the effects of our treatments on participants' decision errors and consumer welfare.

3.4.1 Senders' decisions

This section provides the results for the senders' decisions. We will use these results to compute the winning chances of the lottery tickets for the receivers in Task 2.

In the first part of the experiment the senders had to choose a label for their envelope. 57% of all senders who received an empty envelope labeled it with message A ("This envelope contains \$5."). Envelopes that contained \$5 were automatically labeled with message A.

What does this mean for the receivers? First, with a probability of 0.215 the envelope in Task 1 was sold automatically because it showed message B ("This envelope is empty."). Second, the share of senders that labeled an empty envelope with message A determines the probability that an envelope with message A contained \$5. According to Equation (3.1), an envelope that showed message A contained \$5 with a probability of 0.64.

Our experimental design requires outcome equivalence between Task 1 and the lottery in Task 2. We obtain this outcome equivalence by assigning a winning chance of 0.64 to

the lottery ticket. In addition, the probability that the lottery ticket is sold automatically was set to 0.215. This is identical to the chance that the envelope in Task 1 was labeled with message B and thus sold automatically.

3.4.2 Monotonicity and Carry-Over Effects

This section reports test results for two potential issues that could arise in our experiment and lead to biased estimates of treatment effects. The first issue concerns violations of the monotonic preference assumption while the second issue relates to carry-over effects. Therefore, we report our results on these issues before proceeding with presenting the main results of our experiment in the next section.

We first report the results of our test for violations of monotonic preferences. If subjects violate the monotonic preference assumption, we have to drop them from the main analyses. The reason is that the BDM mechanism that we use in Tasks $1-3$ is only incentive compatible under this assumption (Azrieli, Chambers, and Healy 2018). A subject with strictly monotonic preferences can have at most one switching point from the lottery to the safe outcome. Thus, if subjects chose the safe outcome in a decision problem and the lottery in any of the subsequent decision problems (with strictly higher safe outcome), they violated the strict monotonicity assumption. In our experiment, 35 participants made choices that were inconsistent with this assumption and were thus excluded from the analyses.[8]

Next, we turn to reporting our findings regarding carry-over effects. Carry-over effects occur when a treatment affects the decision of a participant in an unrelated subsequent task (Greenwald 1992, Charness, Gneezy, and Kuhn 2012). In our experiment, this means in particular that the unrelated treatment in Task 1 could have affected decisions in Task 2. In this case, our judgement about a participant's decision error could be distorted and we would, therefore, not be able to perform unbiased analyses of the concerned treatment.

We now report two tests for carry-over effects. First, we regress the decisions in Task 2 on the categorical treatment variable. The treatment should, since it is unrelated to Task 2, not affect decisions in Task 2. We thus expect not to find significant differences in means across the four treatments. The results of this first test are shown in Table 3.2. We see that treatment $T1$ (*correct beta*) had a marginally significant effect on decisions in Task 2 (-0.34 compared to treatment $T0$ and -0.32 compared to treatment $T3$, both $p = 0.071$). Decisions did not significantly differ between treatment $T0$ (baseline), treatment $T2$ (slider), and treatment $T3$ (full info). Although the joint F-test does not reject the *null* ($p = 0.212$), it seems possible that treatment $T1$ affected decisions in

[8]Note that our results would not change qualitatively if they were included in the analyses.

Table 3.2: Effects of treatment in Task 1 on decisions in Task 2

Task 2 Decision	Mean	N	Treatment 0	Treatment 1	Treatment 2
Treatment 0	3.26	93			
Treatment 1	2.97	95	$p = 0.071$		
Treatment 2	3.10	91	$p = 0.349$	$p = 0.277$	
Treatment 3	3.22	86	$p = 0.902$	$p = 0.071$	$p = 0.384$
Whole sample	3.13	365		joint F-test: $p = 0.212$	

Notes. Table shows mean answers in Task 2 by treatment in Task 1. p-values are based on t-tests. N denotes sample size.

Table 3.3: Share of participants whose decisions in Tasks 1 and 2 coincide

Tasks 1 and 2	Treatment 0	Treatment 1	Treatment 2	Treatment 3
Decisions coincide	32%	42%	34%	36%
Decisions differ	68%	58%	66%	64%
N	93	95	91	86

$p = 0.532$, Fisher's exact test

Task 2.[9] Such a potential carry-over effect from treatment $T1$ to answers in Task 2 might distort our measure of the decision error. Therefore, treatment effect estimates for treatment $T1$ have to be interpreted with caution.

Second, we test whether our treatments affected the share of participants who made two identical decisions in Tasks 1 and 2 due to a "taste for consistency" (Falk and Zimmermann 2012). The concern could be that due to a treatment in Task 1, subjects might have adjusted their decision because they recognized that Tasks 1 and 2 were identical in terms of outcomes. Table 3.3 reports the share of decisions that coincide, respectively differ, by treatment status. We see that these shares are not significantly affected by the treatments. Neither the joint test ($p = 0.532$, Fisher's exact test) nor *Wald* tests that compare all treatments pairwise are significant at conventional significance levels. We thus do not find evidence for a carry-over effect of the treatments on the share of coinciding decisions in Tasks 1 and 2.

What do these results mean for our main analyses? First, we will disregard 35 participants from the analyses because of violations of the monotonicity assumption. Second, we will interpret results for treatment $T1$ with great caution because the effects of this treatment might have carried over to decisions in Task 2.

3.4.3 Decision Error and Welfare

In this section we present our main results. Table 3.5 shows how each treatment affected the participants' decision errors. Recall that the decision error is defined as the absolute

[9]Results of our balance checks do not suggest that a randomization issue caused the observed differences.

Table 3.4: Decision errors by treatment and gender

Decision error	All (1)	Men (2)	Women (3)
Treatment 0 (*baseline*)	0.766	0.852	0.689
	(0.917)	(1.115)	(0.697)
Treatment 1 (*beta*)	0.597	0.646	0.551
	(0.807)	(0.823)	(0.797)
Treatment 2 (*slider*)	0.614	0.489	0.789
	(0.650)	(0.559)	(0.732)
Treatment 3 (*full info*)	0.570	0.491	0.653
	(0.708)	(0.576)	(0.823)
Number of participants	365	187	178

Notes. Averages with standard deviations in parentheses.

difference between a subject's decisions in Tasks 1 and 2. The data are thus censored from below at 0. The estimates are based on a simple Tobit model with the decision error as the dependent variable and a categorical treatment indicator as the independent variable. The p-values in this section are calculated using two-sided t-tests, if not otherwise indicated.

When the decisions of men and women are analyzed jointly, we do not see significant effects of our treatments. Even though all three treatments reduced the decision error, the point estimates are not significant ($T1$: -0.177, $p = 0.108$; $T2$: -0.121, $p = 0.281$; $T3$: -0.162, $p = 0.149$).

Analyses of the treatment effects on men and women separately, however, reveal surprising findings. First, women's decision error does not seem to decrease due to our treatments. Women in treatment $T3$ who received full information, i.e. information about the exact share of senders who lied as well as a tool to map this share into the posterior (thus reducing the problem to a simple lottery), did not make significantly better decisions than women who did not receive any additional information (-0.043, $p = 0.781$). The point estimates for the effects of treatments $T1$ (*correct beta*) and $T2$ (*slider*) are insignificant as well ($T1$: -0.140, $p = 0.334$; $T2$: 0.092, $p = 0.587$).

Second, men could improve their decisions in the full information treatment $T3$ (-0.287, $p = 0.075$). The average gap between decisions in Tasks 1 and 2 decreased by 34% relative to the gap in the baseline treatment $T0$. It seems that this improvement was largely driven by the effect of the slider, which, on its own, had an effect of almost identical size on the decision error (treatment $T2$: -0.298, $p = 0.055$). Treatment $T1$ did

not reduce men's decision error significantly.[10]

When we compare the treatment effects on men and women, we see that the slider has a significantly different effect on the two genders. Men benefit more from this treatment than women (0.390, $p = 0.079$, difference-in-differences). The effect of treatment $T3$ (full info) is not significantly different for men and women (0.239, $p = 0.298$, difference-in-differences).

The implications for expected consumer welfare are shown in Table 3.6. In the baseline treatment $T0$, participants left an average consumer surplus of \$0.142 on the table. The decreases in foregone consumer surplus due to treatments $T1$ to $T3$ are not significant in the pooled sample ($T1$: \$−0.041, $p = 0.131$; $T2$: \$−0.040, $p = 0.133$; $T3$: \$−0.042, $p = 0.124$).

In the separate analyses for both genders, we see that men were able to significantly decrease foregone consumer surplus by \$0.095 ($p = 0.025$) in treatment $T2$ and by \$0.091 ($p = 0.038$) in treatment $T3$. These are each reductions of almost half of foregone consumer surplus compared to men's baseline level in treatment $T0$.[11] The reduction of \$0.075 ($p = 0.098$) in treatment $T1$ has to be interpreted with a grain of salt due to the observed carry-over effects. Women, in contrast, do not seem to benefit from the treatments, all three estimates are insignificant ($T1$: \$−0.012, $p = 0.697$; $T2$: \$0.016, $p = 0.638$; $T3$: \$0.004, $p = 0.894$). The availability of the slider in treatment $T2$ helped men to reduce foregone consumer surplus by \$0.109 more than women ($p = 0.049$). We do not see a significant gender difference in the effect of treatments $T1$ (\$0.052, $p = 0.405$) or $T3$ (\$0.090, $p = 0.118$). The results on consumer surplus thus mirror the ones on decision errors and show that the mistakes that subjects made directly translated into losses in consumer surplus.

[10] Remember that treatment $T1$ is potentially compromised due to carry-over effects.

[11] Note that men's foregone consumer surplus in the baseline treatment $T0$ is slightly larger in our sample than women's foregone consumer surplus. The difference is, however, not statistically significant ($p = 0.165$).

Table 3.5: Unconditional marginal effects on decision error

Decision error	All (1)	Men (2)	Women (3)	Gender difference (4)
Treatment 1 (*beta*)	-0.177	-0.221	-0.140	0.065
	(0.110)	(0.165)	(0.145)	(0.169)
Treatment 2 (*slider*)	-0.121	-0.298*	0.092	0.390*
	(0.113)	(0.155)	(0.168)	(0.223)
Treatment 3 (*full info*)	-0.162	-0.287*	-0.043	0.239
	(0.112)	(0.161)	(0.156)	(0.229)
Baseline	0.766***	0.852***	0.689***	-0.153
Treatment 0	(0.095)	(0.163)	(0.097)	(0.178)
N	365	187	178	

Notes. This table reports unconditional average marginal treatment effects on decision error. Estimates are based on a Tobit regression model, data are censored from below at 0. Dependent variable is *decision error*, which is the absolute difference between the answers of a participant in Task 1 and Task 2. Standard errors in parentheses. Column *Gender difference* reports difference-in-differences estimates for gender differences in treatment effects. *Baseline* reports mean decision errors in treatment $T0$ with bootstrapped standard errors in columns (1) to (3).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Unconditional marginal effects on foregone consumer surplus

Foregone Consumer Surplus	All (1)	Men (2)	Women (3)	Gender difference (4)
Treatment 1 (*beta*)	-0.041	-0.075*	-0.012	0.052
	(0.027)	(0.045)	(0.030)	(0.062)
Treatment 2 (*slider*)	-0.040	-0.095**	0.016	0.109**
	(0.027)	(0.042)	(0.034)	(0.055)
Treatment 3 (*full info*)	-0.042	-0.091**	0.004	0.090
	(0.027)	(0.044)	(0.032)	(0.058)
Baseline	0.142***	0.194***	0.095***	-0.069
Treatment 0	(0.034)	(0.068)	(0.019)	(0.050)
N	365	187	178	

Notes. This table reports unconditional average marginal treatment effects on expected foregone consumer surplus. Estimates are based on a Tobit regression model, data are censored from below at 0. Dependent variable is *foregone consumer surplus*, as calculated in equation 3.3. Standard errors in parentheses. Column *Gender difference* reports difference-in-differences estimates for gender differences in treatment effects. *Baseline* reports mean foregone consumer surplus in treatment $T0$ with bootstrapped standard errors in columns (1) to (3).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.5 Conclusion

In this paper, we study whether and how two different informational treatments affect the decision error of men and women in a simple signaling game. The first treatment is designed to alleviate potential negative effects of misspecified beliefs, compared to a baseline treatment. The second treatment reduces the decision error due to incorrect Bayesian updating. Finally, the third treatment combines all elements of the former two.

We find evidence for gender differences in the effects of our treatments. While men's decision quality is significantly improved by a "slider" that maps a prior belief into a posterior probability, we do not find this treatment to affect women's decision errors. Additional information about the actual share of senders who do not send the true message did not lead to a further decrease in decision error for neither gender. Welfare analyses show a substantial decrease in men's foregone consumer surplus due to the slider. Their foregone consumer surplus of \$0.194 was almost halved by our treatment.

These are interesting findings for policy makers in the ongoing discussion about regulating online platforms. In light of the pervasiveness of fake reviews, awareness policies have been discussed as a concrete measure for alleviating the negative effects they have on consumers.[12] Estimates about the share of fake reviews as provided by websites like Fakespot.com are one natural way of providing this information. However, given the high cost involved, obliging all platforms to provide such estimates does not seem practicable. Our findings suggest that providing a relatively simple online tool that allows consumers to correctly map their prior beliefs about the share of fake reviews into the posterior review already improves their purchasing decisions.

Another implication of our findings is that such policies do not have the same effect on women as they have on men. In particular, this study suggests that while such interventions might be an effective consumer protection policy in male-dominated markets, they may not be very effective in female-dominated markets.

[12] On a European level such measures have been brought forward by the European Commissioner for justice, who insists that businesses provide visible information about the reliability of reviews.

3.A Instructions

The sender instructions below do not correspond exactly to what is described in section 2.1. They include an additional task that is unrelated to the experiment itself and we want to briefly explain why we included it.[13] In order to have access to a representative sample *Prolific* requires a minimum of 300 participants per study. Given that our experiment consisted of two studies (the sender part and the receiver part), we had to include 300 subjects in the sender part. To keep the experiment simple and the instructions clear, we matched every sender with two receivers. One for whom Task 1 or 2 and thus the envelope was relevant, and one for whom Task 3 or 4 was relevant. That way, we ensured that each sender sent his envelope to *one* receiver whose decision affected the sender's payoff. It was therefore important to have twice as many receivers as senders.

For budgetary reasons, the only way to conduct our study with a representative sender sample and to maximise the number of receivers while still maintaining a one to two matching, was to recruit 300 sender of which we matched only 200. Of those 300 senders, 100 got assigned an empty envelope, 100 one that contained $5, and 100 were assigned to the additional task.

Note also that the depicted practice exercise below (Task 4 - Part 2) does not show the actual Raven matrix that we used in the experiment.

[13]The task consisted of the same *Cognitive Reflection Test* that was part of receivers' Task 4.

Welcome!

In today's experiment you will be asked to complete three tasks. This will take no longer than 10 minutes and your decisions in these tasks will remain anonymous. At the end of the experiment, you will receive a participation fee of $2 for having completed the three tasks. In addition, we will randomly select one of the three tasks to be relevant. Before each task we will describe in detail how your payment is determined within that task.

Your total earnings from the experiment are the sum of your participation fee and your bonus payment from the randomly selected task.

Consent form and data protection information

The University of Vienna, Universitaetsring 1, 1010 Vienna, Austria is conducting a scientific study using Prolific today.

Your responses will be recorded on our server. The data generated in the study will be separated from the data in the Prolific system after the payment has been completed. The final dataset will not allow any inference to be drawn about the individual participants or the responses of those participants that can be linked to their identity in any way. Correspondingly, the analysis and presentation of all results of this experiment will be anonymized. The anonymous research data will be archived and will possibly be made available to other scientists for further use.

We want to ask for your consent to participate in the experiment and to the implied processing of data. You can revoke your consent at any time, for example, by sending an e-mail to dominik.stelzeneder@univie.ac.at. Your consent is at the same time the legal basis for the processing of data.

For any general information about the processing of data and data protection, especially about your rights to disclosure, correction, deletion, limitations of processing of data, objection, data transmission and complaints to a supervisory body, please contact the data protection officer of University of Vienna: verarbeitungsverzeichnis@univie.ac.at. You also have the right of complaint to the Austrian Data Protection Agency (dsb@dsb.gv.at).

Participation in today's study is entirely voluntary. You have the possibility to exit the study at any time.

You cannot participate in this experiment without consenting.

I have read the data protection information and I consent to participate in the experiment and to the described data processing procedures:

- ○ Yes
- ○ No

Tasks 1 and 2

For this experiment we prepared 200 envelopes in the following way: We put $5 in 100 envelopes. The remaining 100 envelopes remain empty, they thus contain $0.

There are two distinct groups of participants in this experiment, *senders* and *receivers*. You are one of the *senders* (and this will not change throughout the experiment). Each *sender* receives one randomly selected envelope and they are asked to label it with one of two messages:

A) "This envelope contains $5"
B) "This envelope is empty "

The rules are the following:
If the envelope contains $5, you can only choose message A. If, however, the envelope is empty, you can choose freely one of the two messages.

Subsequently, the envelope will be sent to a randomly selected *receiver*. For the *receiver* the experiment will proceed as follows. Each *receiver* will have to name the lowest offer they are willing to accept in exchange for the envelope – the *minimum selling price*. The computer will then generate a random amount between $0 and $5 and offer it to the *receiver*.

If the envelope is labelled with message A ("This envelope contains $5"), the *receiver* either opens or sells the envelope:

- If the computer's offer is below the *minimum selling price*, the *receiver* opens the envelope. In this case the *receiver* receives whatever amount the envelope contains and you get $3.
- If the computer's offer is higher than or as high as the *minimum selling price*, the *receiver* sells the unopened envelope. In this case the *receiver* receives the offered amount and you get $0.

If the envelope is labelled with message B ("This envelope is empty"), the *receiver* sells the envelope irrespective of the submitted *minimum selling price*.

- In this case the *receiver* receives the offered amount and you get $0.

These instructions are provided to all participants of the experiment.

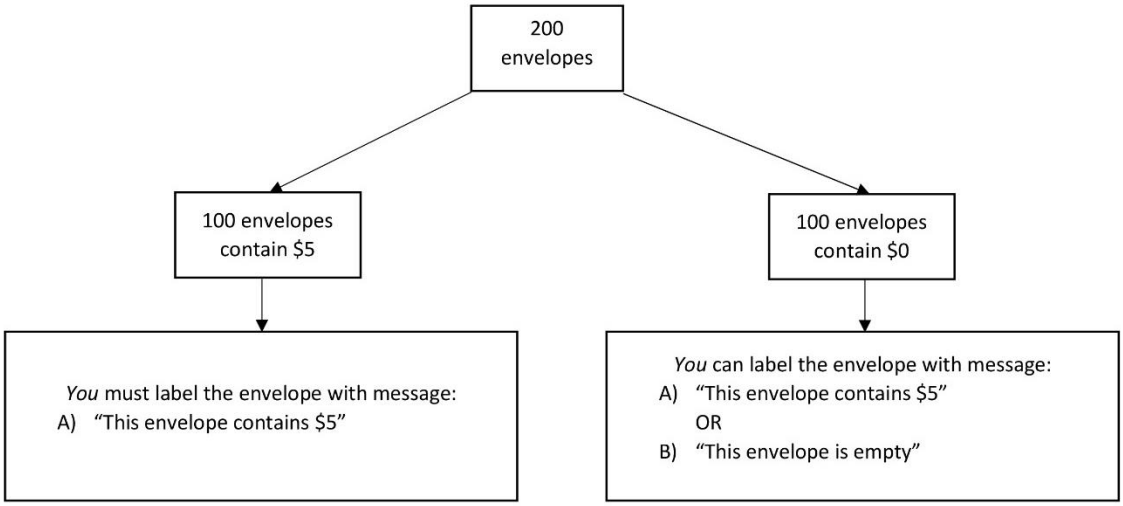
Your task will be to label your envelope.

If your letter contains $5, it will automatically be labelled with message A ("This envelope contains $5").

If it is empty, you will be asked to answer the following question:

Which message do you choose to label your envelope with?

Tasks 1 and 2



If *you* label the envelope with message A ("This envelope contains $5"):

The offer is below the *receiver's* minimum selling price:
The *receiver* gets the content of the envelope ($0 or $5) and *you* get $3.

OR

The offer is higher than or as high as the *receiver's* minimum selling price:
The *receiver* gets the amount offered (between $0 and $5) and *you* get $0.

If *you* label the envelope with message B ("This envelope is empty"):

The *receiver* gets the amount offered (between $0 and $5) and *you* get $0.

In Task 1 on the next page you will answer the following question:

Which message do you choose to label your envelope with?

Task 1 - Decision

By clicking on the ‚+' symbols you can access the instructions again. Please take your time and consider your decision carefully.

+ Information on Task 1

+ Graphical Illustration

Your envelope is empty.

Please answer the following question:

Which message do you choose to label your envelope with?

Message A: "This envelope contains $5"
Message B: "This envelope is empty"

Please click 'Next' after you have selected a message.

Task 2 - Decision

Your envelope contains 5 $.

Your envelope will therefore automatically be labelled with message A (“This envelope contains $5”).
Please click ‘Next’ to proceed.

Task 3

Please answer the following questions. You will receive $0.40 for each correct answer if Task 3 is randomly selected to be relevant.

Question 1:
A bat and a ball together cost $1.10. The bat is $1 more expensive than the ball. How much does the ball cost?

Answer (in cents):

Question 2:
5 machines take 5 minutes to produce 5 items. How long do 100 machines take to produce 100 items?

Answer (in minutes):

Question 3:
A lake is partially covered with lily pads. The area that is covered by lily pads doubles every day. If it takes 48 days for the lily pads to cover the whole lake, how long does it take until half of the lake is covered?

Answer (in days):

Final Questions

What is your gender?

- Woman
- Man
- Non-binary
- Prefer not to say

How old are you?

What is your highest completed level of education?

Thank you for your participation!

You earn $2 for partaking in today's experiment. Task 2 was randomly selected to be relevant, you thus received an envelope that contained $5.

Hence, you automatically labelled it with message A ("This envelope contains $5"). Therefore, your additional earnings depend on the *receiver*'s decision. If the receiver opens the envelope, you will get a bonus of $3. Otherwise, you will not get a bonus. Hence, you will receive either $2 or $5 in total. We will send you your participation fee within two days. As soon as the *receiver* of your envelope has made a decision and your bonus for Task 2 has been determined, we will send you your bonus. This should not take longer than a week.

Feedback

Lastly, we kindly ask for your feedback on this study. Was anything unclear? What do you think could be improved?

Welcome!

In today's experiment you will solve 4 Tasks. Each task will take less than 15 minutes and your decisions will remain anonymous. At the end of the experiment, you will receive a participation fee of $ 4 for having completed the four tasks. In addition, we will randomly select one of the tasks and pay you based on your decisions or your performance in that task. In addition, there will be two additional tasks, that will contribute to your earnings. Before each task we describe in detail how your payment is determined within that task. At the end of the experiment, we inform you about which task was selected, and we break down your total earnings. They will represent the sum of your participation fee, your earnings in the randomly selected task, and your earnings in the additional tasks.

Consent form and data protection information

The University of Vienna, Universitaetsring 1, 1010 Vienna, Austria is conducting a scientific study using Prolific today.

Your responses will be recorded on our server. The data generated in the study will be separated from the data in the Prolific system after the payment has been completed. The final dataset will not allow any inference to be drawn about the individual participants or the responses of those participants that can be linked to their identity in any way. Correspondingly, the analysis and presentation of all results of this experiment will be anonymized. The anonymous research data will be archived and will possibly be made available to other scientists for further use.

We want to ask for your consent to participate in the experiment and to the implied processing of data. You can revoke your consent at any time, for example, by sending an e-mail to dominik.stelzeneder@univie.ac.at. Your consent is at the same time the legal basis for the processing of data.

For any general information about the processing of data and data protection, especially about your rights to disclosure, correction, deletion, limitations of processing of data, objection, data transmission and complaints to a supervisory body, please contact the data protection officer of University of Vienna: verarbeitungsverzeichnis@univie.ac.at. You also have the right of complaint to the Austrian Data Protection Agency (dsb@dsb.gv.at).

Participation in today's study is entirely voluntary. You have the possibility to exit the study at any time.

You cannot participate in this experiment without consenting.

I have read the data protection information and I consent to participate in the experiment and to the described data processing procedures:

- ○ Yes
- ○ No

Task 1

For this experiment we prepared 200 envelopes in the following way: We put $5 in 100 envelopes. The remaining 100 envelopes remain empty, they thus contain $0.

There are two distinct groups of participants in this experiment, *senders* and *receivers*. Each *sender* receives one randomly selected envelope, and they are asked to label it with one of two messages:

A) "This envelope contains $5"
B) "This envelope is empty"

The rules are the following: if the envelope contains $ 5, the *sender* can only choose message A. If, however, the envelope is empty, the *sender* can choose freely one of the two messages.

Subsequently, the envelope will be sent to a randomly selected *receiver*.

You are one of the *receivers*. You will have to name the lowest offer that you are willing to accept in exchange for the envelope – the *minimum selling price*. The computer will then generate a random amount between $0 and $5 and offer it to you.

If the envelope is labelled with message A ("This envelope contains $5"), you either open or sell the envelope:

- If the computer's offer is below the *minimum selling price*, you open the envelope. In this case you receive whatever amount the envelope contains, and the *sender* gets $3.
- If the computer's offer is higher than or as high as the *minimum selling price*, you sell the unopened envelope. In this case you receive the offered amount, and the *sender* gets $0.

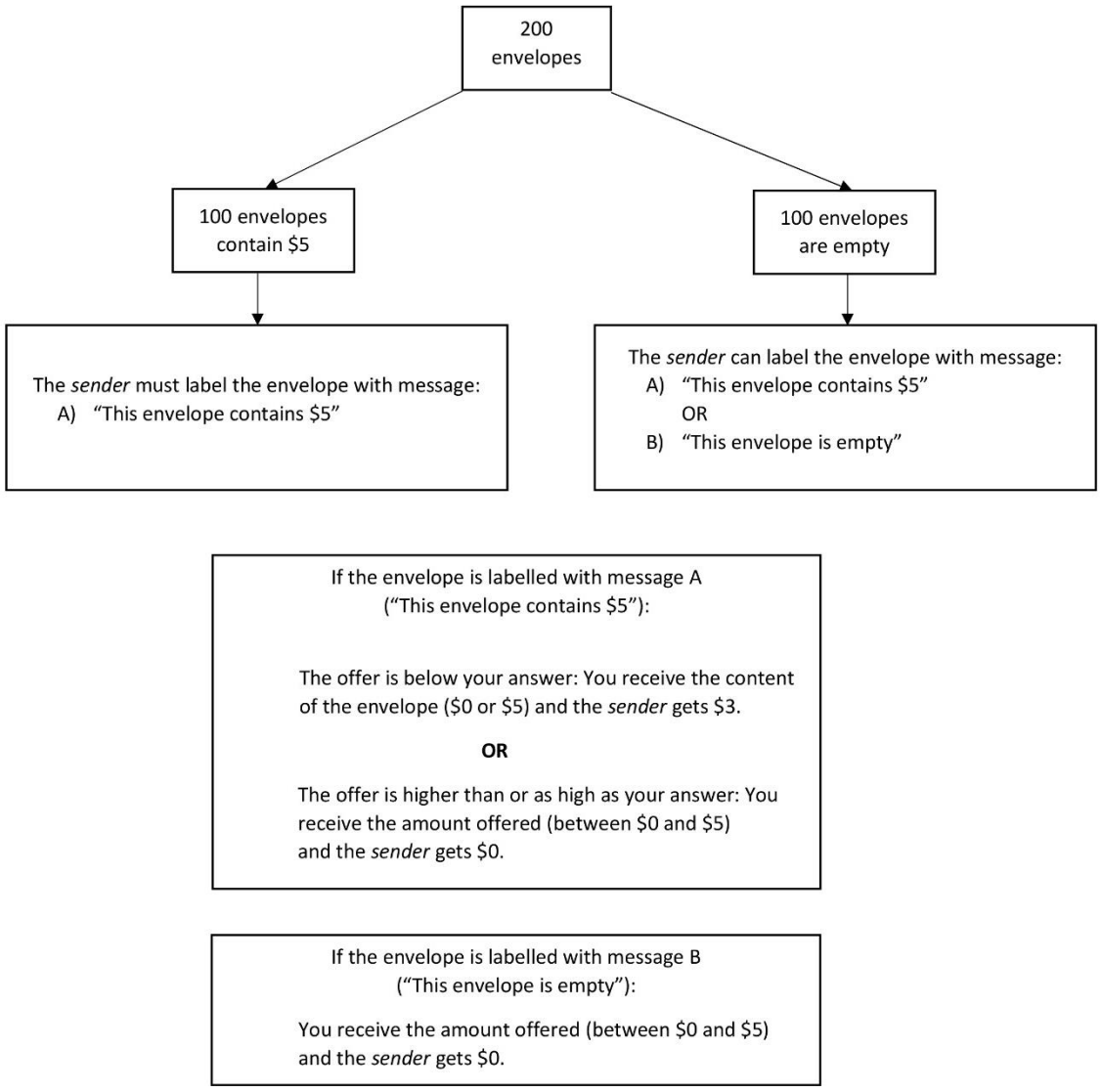
If the envelope is labelled with message B ("This envelope is empty"), you sell the envelope irrespective of the submitted *minimum selling price*.

- In this case you receive the offered amount, and the *sender* gets $0.

These instructions are provided to all participants of the experiment.

In a moment, we will ask you to submit a *minimum selling price*. The illustration on the next page summarizes the instructions above.

Task 1



The question you will be asked is:

If your envelope is labeled with message A ("This envelope contains $5"), for how much money would you be willing to sell it to us?

Comprehension Questions

Before we turn to Task 1, please answer the following comprehension questions. We will show you the correct solutions on the next page.

Question 1

How many of 200 envelopes contain $5?

Answer:

Question 2

How many of 200 envelopes are empty?

Answer:

Question 3

If the envelope is labeled with the message "This envelope contains $5", then it contains $5 with certainty.

This is:

Question 4

How much does the *sender* earn if you open the envelope?

Answer (in $):

Question 5

How much does the *sender* earn if you do not open the envelope?

Answer:

Question 6

If your envelope is labeled with the message "This envelope contains $5", then you will get the amount offered to you if the offer is higher than or as high as your minimum selling price.

This is:

Solutions

This page shows you the correct answers for the comprehension questions. Below, you can see whether your answer was correct and you can access additional information by clicking on „+ more info“.

Question 1

How many of the 200 envelopes contain $5? You answered: 100 envelopes
Your answer is correct

+ more info
„For this task we have prepared 200 envelopes in the following way: We put $5 in 100 envelopes. The remaining 100 envelopes remain empty, they thus contain $0."

Question 2

How many of the 200 envelopes are empty? You answered: 100 envelopes
Your answer is correct.

+ more info
„For this task we have prepared 200 envelopes in the following way: We put $5 in 100 envelopes. The remaining 100 envelopes remain empty, they thus contain $0.“

Question 3

If the envelope is labelled with message A ("This envelope contains $5"), then it contains $5 with certainty. You answered: *„False“*
The statement is false. Therefore, your answer is correct.

+ more info

"Each *sender* receives one randomly selected envelope and they are asked to label it with one of the two messages:

A. "This envelope contains $5"
B. "This envelope is empty"

The rules are the following: if the envelope contains $5, the *sender* can only choose message A. If, however, the envelope is empty, the *sender* can choose freely one of the two messages."

Question 4

How much does the *sender* earn if you open the envelope? You answered: $3
Your answer is correct

+ more info

"If the computer's offer is below the *minimum selling price*, you open the envelope. In this case you receive whatever amount the envelope contains, and the *sender* gets $3."

Question 5

How much does the *sender* earn if you do not open the envelope? You answered: $0
Your answer is correct

+ more info

"If the computer's offer is higher than or as high as the *minimum selling price*, you sell the unopened envelope. In this case you receive the offered amount, and the *sender* gets $0."

Question 6

If your envelope is labelled with message A ("This envelope contains $5"), then you get the amount offered to you if the offer is higher than or as high as your *minimum selling price*.
You answered: *„True“*
The statement is true. Therefore, your answer is correct.

+ more info

"If the envelope is labelled with message A ("This envelope contains $5"), you either open or sell the envelope:

- If the computer's offer is below the *minimum selling price*, you open the envelope. In this case you receive whatever amount the envelope contains, and the *sender* gets $3.
- If the computer's offer is higher than or as high as the *minimum selling price*, you sell the unopened envelope. In this case you receive the offered amount, and the *sender* gets $0."

Task 1

By clicking on the ‚+' symbols you can access the instructions again. Please take your time and consider your decision carefully.

+ Information on Task 1

+ Graphical Illustration

Additional information:

75% of all senders whose envelopes were empty, labeled them with the message A ("This envelope contains $5").

Below you can use a calculator that was designed specifically for this task. You can submit an estimate of the share of *senders* who labeled their empty envelopes with message A ("This envelope contains $5"). Based on your estimate, the calculator computes the probability that *your* envelope contains either $0 or $5. You can submit multiple estimates and use the calculator as often as you wish.

Input:

If % of those *senders* whose envelopes were empty, labeled them with the message A ("This envelope contains $5") and you receive an envelope that is labeled with message A, then:

- Your envelope contains $5 with a probability of %.
- Your envelope is empty with a probability of %.

Please answer the following question:

If your envelope is labeled with the message A ("This envelope contains $5"), for how much money would you be willing to sell it to us?

The amount can include up to two decimals for cents.

Your answer: $0.4

If we offer you between $0 and $0.39, you will open the envelope.

If we offer you $0.40 or more, you will sell the envelope.

Task 2

In this task you will receive a lottery ticket. The conditions of the lottery are explained in detail below. We will then ask you to name the lowest offer that you are willing to accept in exchange for the lottery ticket – the minimum selling price. The computer will then generate a random amount between $0 and $5 and offer it to you.
If the offer is below the minimum selling price, you will keep your ticket and participate in the lottery.
If the offer is higher than or as high as the minimum selling price, you will sell the lottery ticket in exchange for the offered amount.
However, with a probability of 12,5% you automatically sell the lottery ticket regardless of your answer.

The lottery in Task 2 can lead to earning for you and for one of the *senders* from Task 1. The conditions of the lottery are as follows:

Lottery

- With a probability of 57% you earn $5 and one *sender* earns $5. The *sender* in this case will be chosen randomly among all those whose envelope contained $5.

- With a probability of 42% you earn $0 and one *sender* earns $5. The *sender* in this case will be chosen randomly among all those whose envelope was empty and who chose to label it with the following message: "This envelope contains $5"

Please answer the following question:

For how much money would you be willing to sell us your lottery ticket?

The amount can include up to two decimals for cents.

Your answer: $

If we offer you between $0.00 and $, you will participate in the lottery.

If we offer you $ or more, you will sell the lottery ticket.

Task 3

In this task you will receive a lottery ticket. The conditions of the lottery are the same as in Task 2, except that now the *senders* do not have the chance to get a payment.
We will ask you to name the lowest offer that you are willing to accept in exchange for the lottery ticket – the minimum selling price. The computer will then generate a random amount between $0 and $5 and offer it to you.
If the offer is below the minimum selling price, you will keep your ticket and participate in the lottery.
If the offer is higher than or as high as the minimum selling price, you will sell the lottery ticket in exchange for the offered amount.

However, with a probability of 0,125% you automatically sell the lottery ticket regardless of your answer.

The conditions of the lottery are as follows:

Lottery

- With a probability of 57% you earn $5.

- With a probability of 42% you earn $0.

Please answer the following question:

For how much money would you be willing to sell us your lottery ticket?

The amount can include up to two decimals for cents.

Your answer: $

If we offer you between $0.00 and $ you will participate in the lottery.

If we offer you $ or more, you will sell the lottery ticket.

Task 4

This task consists of 2 parts. In the first part you must answer 3 questions and you earn $0.50 per correct answer.

Click "Next" to start with the first part of Task 4.

Task 4 - Part 1

Question 1:
A bat and a ball together cost $1.10. The bat is $1 more expensive than the ball. How much does the ball cost?

Answer (in cents):

Question 2:
5 machines take 5 minutes to produce 5 items. How long do 100 machines take to produce 100 items?

Answer (in minutes):

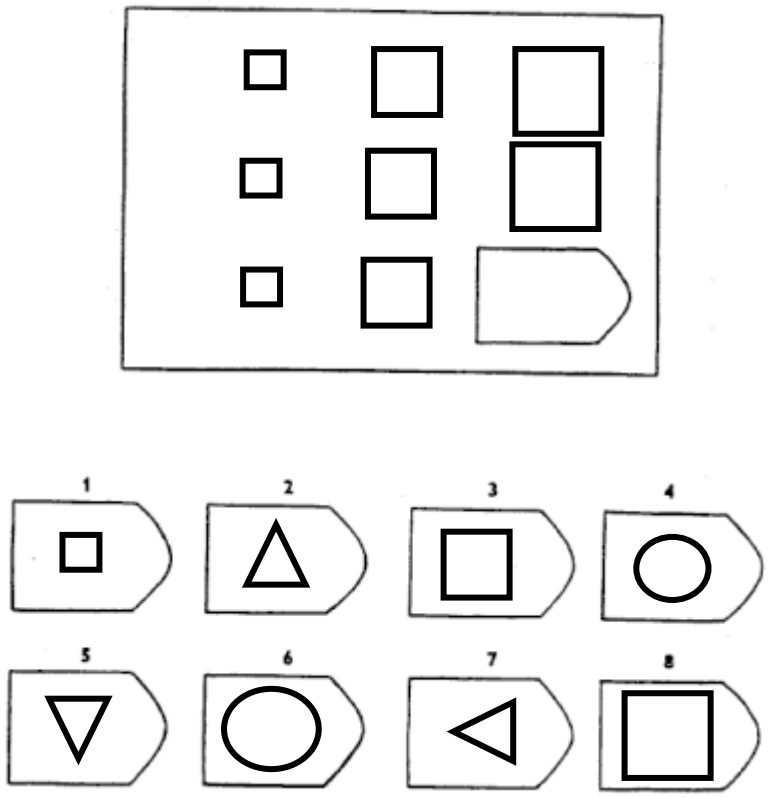
Question 3:
A lake is partially covered with lily pads. The area that is covered by lily pads doubles every day. If it takes 48 days for the lily pads to cover the whole lake, how long does it take until half of the lake is covered?

Answer (in days):

Task 4 - Part 2

In the second part of Task 4, we will show you a series of matrices containing patterns but with one missing piece. Your task is to pick the pattern that completes the matrix. For each matrix there is exactly one correct solution.

Here is an example:



Which one of these patterns completes the image?

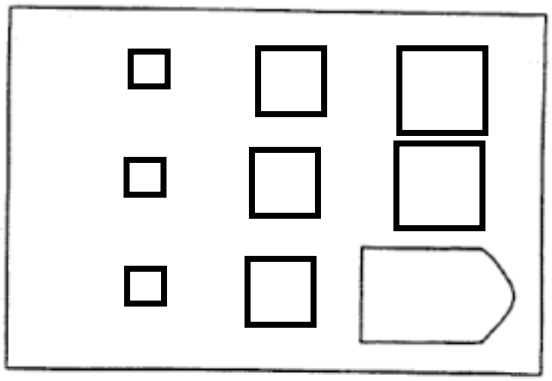
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

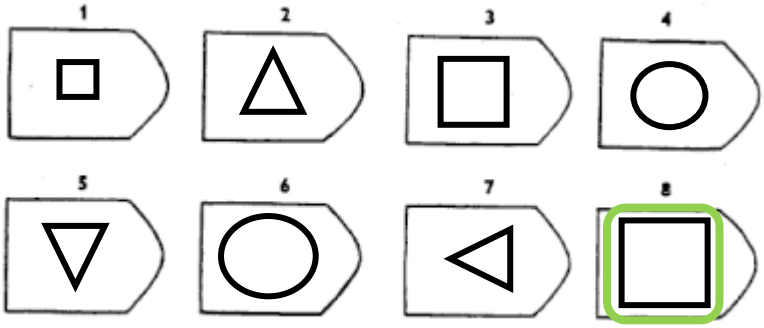
Please select one option and click on the "Next" button. You will see the correct solution on the next page.

Task 4 - Part 2

Your answer was correct.
Total number of correct answers: 1

The correct answer for the introductory example is pattern number 8.





Which one of these patterns completes the image?

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

There are 34 matrices, and you have three minutes to solve as many as possible. You will earn $0.30 per correct answer. If Task 4 is chosen as the relevant one for your payoff, your earnings from part one and part two will be added and paid to you.

The problems are easy at the beginning, and gradually become more difficult. The timer will start when you click “Next”.

Additional Task A1

Please think back to Task 1, in which *senders* were asked to label an envelope with one of these two messages:

A. “This envelope contains $5”
B. “This envelope is empty”

100 senders had an envelope that was empty. How many of these 100 senders do you think chose message A?
Answer:

You will receive $1 if your answer deviates by less than three from the correct answer.

Additional Task A2

For each of the following rows, please choose either A or B. One of the rows will be randomly selected for payment at the end of the experiment.

Row		Option A	Option B	
1	○	100% chance of winning $0.05	75% chance of winning $1 25% chance of winning $0	○
2	○	100% chance of winning $0.10	75% chance of winning $1 25% chance of winning $0	○
3	○	100% chance of winning $0.15	75% chance of winning $1 25% chance of winning $0	○
4	○	100% chance of winning $0.20	75% chance of winning $1 25% chance of winning $0	○
5	○	100% chance of winning $0.25	75% chance of winning $1 25% chance of winning $0	○
6	○	100% chance of winning $0.30	75% chance of winning $1 25% chance of winning $0	○
7	○	100% chance of winning $0.35	75% chance of winning $1 25% chance of winning $0	○
8	○	100% chance of winning $0.40	75% chance of winning $1 25% chance of winning $0	○
9	○	100% chance of winning $0.45	75% chance of winning $1 25% chance of winning $0	○
10	○	100% chance of winning $0.50	75% chance of winning $1 25% chance of winning $0	○
11	○	100% chance of winning $0.55	75% chance of winning $1 25% chance of winning $0	○

12	○	100% chance of winning $0.60	75% chance of winning $1 25% chance of winning $0	○
13	○	100% chance of winning $0.65	75% chance of winning $1 25% chance of winning $0	○
14	○	100% chance of winning $0.70	75% chance of winning $1 25% chance of winning $0	○
15	○	100% chance of winning $0.75	75% chance of winning $1 25% chance of winning $0	○
16	○	100% chance of winning $0.80	75% chance of winning $1 25% chance of winning $0	○
17	○	100% chance of winning $0.85	75% chance of winning $1 25% chance of winning $0	○
18	○	100% chance of winning $0.90	75% chance of winning $1 25% chance of winning $0	○
19	○	100% chance of winning $0.95	75% chance of winning $1 25% chance of winning $0	○
20	○	100% chance of winning $1.00	75% chance of winning $1 25% chance of winning $0	○

Thank you for participating in today's experiment!

Before informing you about your earnings and the payment modalities we would like you to answer a few brief questions.

Final Questions

What is your gender?

- ○ Man
- ◉ Woman
- ○ Non-binary
- ○ Prefer not to say

What is your age?

What is your highest level of education?

Your Earnings

We will now break down your total earning:

1. For participating in the experiment, you earn $4.00.

2. Task 2 was randomly selected to be relevant. In Task 2 you stated that you were willing to sell your lottery ticket for $3.0 or more. The random device generated an offer of $0.95 for you.

   With a probability of 12.5% you automatically sell the lottery ticket. Random device 1 determined that this is not the case. The offer is therefore compared to your minimum selling price. Since it is lower than your minimum selling price, you play the lottery with a winning probability of 57%. Random device 2 determined that your lottery ticket did not win. You earn $0. A randomly chosen *sender* that labelled his or her empty envelope with message A ("This envelope contains $5") earns $5.

3. In Additional Task A1 you answered that 58 senders of an empty envelope labelled it with message A ("This envelope contains $5"). The correct answer is 57. Your answer deviates by more than two from the correct answer, your earnings thus do not increase. / Your answer deviates by less than three from the correct answer, your earnings thus increase by $1.

4. In Additional Task A2, item 1 was chosen. You chose Option A. Your earnings increase by $0.05.

In today's experiment you earned a total of $6.05.

Bibliography

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. "When should you adjust standard errors for clustering?" In: *The Quarterly Journal of Economics* 138.1 (Oct. 2022), pp. 1–35. ISSN: 0033-5533. DOI: 10.1093/qje/qjac038. URL: https://doi.org/10.1093/qje/qjac038.

Aguirregabiria, Victor. "Identification of firms' beliefs in structural models of market competition". In: *Canadian Journal of Economics/Revue canadienne d'économique* 54.1 (2021), pp. 5–33. DOI: 10.1111/caje.12503.

Aichholzer, Julian, Sylvia Kritzinger, Markus Wagner, Nicolai Berk, Hajo Boomgaarden, and Wolfgang C. Müller. *AUTNES Comparative Study of Electoral Systems Post-Election Survey 2017 (CSES Edition)*. 2018. DOI: 10.11587/w193uz.

Aoyagi, Masaki, Guillaume R. Fréchette, and Sevgi Yuksel. "Beliefs in Repeated Games". In: *Working paper* (May 2022).

Atkinson, John W. "Motives in Fantasy, Action, and Society". In: ed. by John W. Atkinson. New York: D. Van Nostrand, 1958. Chap. Towards experimental analysis of human motivation in terms ofmotives, expectancies, and incentives, pp. 288–305.

Au, Pak Hung and Keiichi Kawai. "Competitive disclosure of correlated information". In: *Economic Theory* 72.3 (2021), pp. 767–799. DOI: 10.1007/s00199-018-01171-7.

— "Competitive information disclosure by multiple senders". In: *Games and Economic Behavior* 119 (2020), pp. 56–78. DOI: 10.1016/j.geb.2019.10.002.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy. "Incentives in Experiments: A Theoretical Analysis". In: *Journal of Political Economy* 126.4 (2018), pp. 1472–1503. DOI: 10.1086/698136.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak. "Measuring utility by a single-response sequential method". In: *Behavioral Science* 9.3 (1964), pp. 226–232. DOI: 10.1002/bs.3830090304.

Benjamin, Daniel J. "Handbook of Behavioral Economics - Foundations and Applications 2". In: ed. by Doug Bernheim, Stefano DellaVigna, and David Laibson. Elsevier Press, 2019. Chap. Errors in Probabilistic Reasoning and Judgment Biases. Pp. 69–186.

Bennett, Linda L.M. and Stephen Earl Bennett. "Enduring Gender Differences in Political Interest". In: *American Politics Quarterly* 17.1 (1989), pp. 105–122. DOI: 10.1177/1532673x8901700106.

Berinsky, Adam J. and Gabriel S. Lenz. "Education and Political Participation: Exploring the Causal Link". In: *Political Behavior* 33.3 (2010), pp. 357–373. DOI: 10.1007/s11109-010-9134-9.

Board, Simon and Jay Lu. "Competitive Information Disclosure in Search Markets". In: *Journal of Political Economy* 126.5 (2018), pp. 1965–2010. DOI: 10.1086/699211.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. "Beliefs about Gender". In: *American Economic Review* 109.3 (2019), pp. 739–773. DOI: 10.1257/aer.20170007.

Bourjade, Sylvain and Bruno Jullien. "The roles of reputation and transparency on the behavior of biased experts". In: *The RAND Journal of Economics* 42.3 (2011), pp. 575–594.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. "Bootstrap-based improvements for inference with clustered errors". In: *The Review of Economics and Statistics* 90.3 (Aug. 2008), pp. 414–427.

Carpenter, Jeffrey, Emiliano Huet-Vaughn, Peter Hans Matthews, Andrea Robbett, Dustin Beckett, and Julian Jamison. "Choice Architecture to Improve Financial Decision Making". In: *The Review of Economics and Statistics* 103.1 (2021), pp. 102–118. DOI: 10.1162/rest_a_00881.

Chadwick, Laura and Gary Solon. "Intergenerational Income Mobility Among Daughters". In: *American Economic Review* 92.1 (2002), pp. 335–344.

Chaisemartin, Clément de and Jaime Ramirez-Cuellar. "At What Level Should One Cluster Standard Errors in Paired Experiments, and in Stratified Experiments with Small Strata?" In: *American Economic Journal: Applied Economics* (2023). forthcoming.

Charness, Gary, Uri Gneezy, and Michael A. Kuhn. "Experimental methods: Between-subject and within-subject design". In: *Journal of Economic Behavior & Organization* 81.1 (2012), pp. 1–8. DOI: 10.1016/j.jebo.2011.08.009.

Chen, Daniel L., Martin Schonger, and Chris Wickens. "oTree—An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9 (2016), pp. 88–97.

Ciminelli, Gabriele, Cyrille Schwellnus, and Balazs Stadler. "Sticky floors or glassceilings? The role of humancapital, working timeflexibility and discriminationin the gender wage gap". In: *OECD Economics Department Working Papers* 1668 (2021). OECD Publishing, Paris. DOI: https://doi.org/10.1787/02ef3235-en.

Clarke, Kevin A. "The Phantom Menace: Omitted Variable Bias in Econometric Research". In: *Conflict Management and Peace Science* 22.4 (2005), pp. 341–352. DOI: 10.1080/07388940500339183.

Congiu, Luca, Salvatore Nunnari, and Emiliano Sandri. *Database of Laboratory Experiments in Top Economics Journals*. retrieved on June 27, 2022. 2022. URL: https://docs.google.com/spreadsheets/d/1434ApdJsvdtRNIpDR_O3h-kmUffcL24_cJj7BXyCezs/edit#gid=491277435.

Costa-Gomes, Miguela A. and Georg Weizsäcker. "Stated Beliefs and Play in Normal-Form Games". In: *Review of Economic Studies* 75.3 (July 2008), pp. 729–762.

Crawford, Vincent P. and Joel Sobel. "Strategic Information Transmission". In: *Econometrica* 50.6 (1982), p. 1431. DOI: 10.2307/1913390.

Cárdenas, Juan-Camilo, Anna Dreber, Emma von Essen, and Eva Ranehill. "Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden". In: *Journal of Economic Behavior and Organization* 83 (2012), pp. 11–23.

Croke, Kevin, Guy Grossman, Horacio A. Larreguy, and John Marshall. "Deliberate Disengagement: How Education Can Decrease Political Participation in Electoral Authoritarian Regimes". In: *American Political Science Review* 110.3 (2016), pp. 579–600. DOI: 10.1017/s0003055416000253.

Dang, Thang. "Quasi-experimental evidence on the political impacts of education in Vietnam". In: *Education Economics* 27.2 (2018), pp. 207–221. DOI: 10.1080/09645292.2018.1554101.

Dellavigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao. "Voting to Tell Others". In: *The Review of Economic Studies* 84.1 (2017), pp. 143–181. DOI: 10.1093/restud/rdw056.

Deloitte. *Women in the boardroom. A global perspective*. online. Deloitte Global Boardroom Program. Seventh edition. 2022.

Donald, Stephen G and Kevin Lang. "Inference with Difference-in-Differences and Other Panel Data". In: *Review of Economics and Statistics* 89.2 (2007), pp. 221–233. DOI: 10.1162/rest.89.2.221.

Dostie-Goulet, Eugénie. "Social networks and the development of political interest". In: *Journal of Youth Studies* 12.4 (2009), pp. 405–421. DOI: 10.1080/13676260902866512.

Dreber, Anna, Emma von Essen, and Eva Ranehill. "Outrunning the gender gap-boys and girls compete equally". In: *Experimental Economics* 14 (2011), pp. 567–582.

Dunn, Kris. "Left-Right identification and education in Europe: A contingent relationship". In: *Comparative European Politics* 9.3 (2011), pp. 292–316. DOI: 10.1057/cep.2010.17.

Eckel, Catherine C. and Philip J. Grossman. “Sex differences and statistical stereotyping in attitudes toward financial risk”. In: *Evolution and Human Behavior* 23 (2002), pp. 281–295.

Edwards, Ward, Lawrence D. Phillips, William L. Hays, and Barbara Goodman. “Probabilistic Information Processing Systems: Design and Evaluation”. In: *IEEE Transactions on Systems Science and Cybernetics* 4.3 (Sept. 1968), pp. 248–265.

Ennser-Jedenastik, Laurenz. “A Welfare State for Whom? A Group-based Account of the Austrian Freedom Party's Social Policy Profile”. In: *Swiss Political Science Review* 22.3 (2016), pp. 409–427. DOI: 10.1111/spsr.12218.

Europe Direct and VieCER. *Fragebogen-Entwurf: "Demokratie in Europa"*. Unpublished.

Falk, A. and F. Zimmermann. “A Taste for Consistency and Survey Response Behavior”. In: *CESifo Economic Studies* 59.1 (2012), pp. 181–193. DOI: 10.1093/cesifo/ifs039.

Fréchette, Guillaume R., Kim Sarnoff, and Leeat Yariv. “Experimental Economics: Past and Future”. In: *Annual Review of Economics* 14.1 (2022), pp. 777–794. DOI: 10.1146/annurev-economics-081621-124424.

Frederick, Shane. “Cognitive Reflection and Decision Making”. In: *Journal of Economic Perspectives* 19.4 (2005), pp. 25–42. DOI: 10.1257/089533005775196732.

Friedman, Willa, Michael Kremer, Edward Miguel, and Rebecca Thornton. “Education as Liberation?” In: *Economica* 83.329 (2015), pp. 1–30. DOI: 10.1111/ecca.12168.

Fulghieri, Paolo, Günter Strobl, and Han Xia. “The Economics of Solicited and Unsolicited Credit Ratings”. In: *Review of Financial Studies* 27.2 (Feb. 2014), pp. 484–518. DOI: 10.1093/rfs/hht072.

Funk, Patricia. “Social Incentives and Voter Turnout: Evidence from the Swiss Mail Ballot System”. In: *Journal of the European Economic Association* 8.5 (2010), pp. 1077–1103. DOI: 10.1162/jeea_a_00015.

Gerber, Alan S., Gregory A. Huber, David Doherty, Conor M. Dowling, and Shang E. Ha. “Personality and Political Attitudes: Relationships across Issue Domains and Political Contexts”. In: *American Political Science Review* 104.1 (2010), pp. 111–133. DOI: 10.1017/s0003055410000031.

Gillen, Ben, Erik Snowberg, and Leeat Yarif. “Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study”. In: *Journal of Political Economy* 127.4 (2019), pp. 1826–1863.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. “Performance in Competitive Environments: Gender Differences”. In: *The Quarterly Journal of Economics* 118.3 (2003), pp. 1049–1074.

Gneezy, Uri and Aldo Rustichini. “Gender and Competition at a Young Age”. In: *The American Economic Review* 94.2 (2004), pp. 377–381.

Günther, Christina, Neslihan Arslan Ekinci, Christiane Schwieren, and Martin Strobel. "Women can't jump? - An experiment on competitive attitudes and stereotype threat". In: *Journal of Economic Behavior and Organization* 75 (2010), pp. 395–401.

Green, Donald P., Shang E. Ha, and John G. Bullock. "Enough Already about "Black Box" Experiments: Studying Mediation Is More Difficult than Most Scholars Suppose". In: *The ANNALS of the American Academy of Political and Social Science* 628.1 (2010), pp. 200–208. DOI: 10.1177/0002716209351526.

Greenwald, Anthony G. "Within-subjects designs: To use or not to use?" In: *Methodological issues & strategies in clinical research.* American Psychological Association, 1992, pp. 157–167. DOI: 10.1037/10109-021.

Grosse, Niels Daniel, Gerhard Riener, and Markus Dertwinkel-Kalt. "Explaining gender differences in competitiveness: Testing a theory on gender-task stereotypes". In: *Working Paper* (2014).

Harrison, Glenn, Jimmy Martínez-Correay, Karlijn Morsinkz, Jia Min Ng, and Todd Swarthout. "Compound Risk and the Welfare Consequencesof Insurance". In: *Working Paper* (Dec. 2020).

Harrison, Glenn, Karlijn Morsink, and Mark Schneider. "Do No Harm? The Welfare Consequences of Behavioural Interventions". In: *Working Paper* (July 2020).

He, Sherry, Brett Hollenbeck, and Davide Proserpio. "Exploiting Social Media for Fake Reviews: Evidence from Amazon and Facebook". In: *ACM SIGecom Exchanges* 19.2 (Nov. 2021), pp. 68–74.

Hernandez-Arenaz, Iñigo. "Stereotypes and tournament self-selection: A theoretical and experimental approach". In: *European Economic Review* 126 (2020), p. 103448. DOI: 10.1016/j.euroecorev.2020.103448.

Holt, Charles A. and Susan K. Laury. "Risk Aversion and Incentive Effects". In: *The American Economic Review* 92.5 (2002), pp. 1644–1655.

Hyman, Herbert Hiram. *Education's Lasting Influence on Values*. University of Chicago Press, 1979. ISBN: 0226365425.

Institut für Neue Kulturtechnologien/t0. *Wahlkabine*. https://wahlkabine.at/. Retrieved on January 27, 2022. 2019.

Iriberri, Nagore and Pedro Rey-Biel. "Stereotypes are only a threat when beliefs are reinforced: On the sensitivity of gender differences in performance under competition to information provision". In: *Journal of Economic Behavior and Organization* 135 (2017), pp. 99–111.

Janssen, Maarten C.W. and Santanu Roy. "Regulating Product Communication". In: *American Economic Journal: Microeconomics* 14.1 (2022), pp. 245–283. DOI: 10.1257/mic.20190187.

Jendraszek, Patricia Anne. “Misconceptions of Probability Among Future Teachers of Mathematics”. PhD thesis. Columbia University, 2008.

Karni, Edi. “A Mechanism for Eliciting Probabilities”. In: *Econometrica* 77.2 (2009), pp. 603–606. DOI: 10.3982/ecta7833.

Krupka, Erin L. and Roberto A. Weber. “Identifying Social Norms Using Coordination Games: Why does Dictator Game Sharing Vary?” In: *Journal of the European Economic Association* 11.3 (2013), pp. 495–524. DOI: 10.1111/jeea.12006.

Lacasse, Hugo. *Bunte baum image*. ID: 9934554. 2023. URL: https://de.123rf.com/photo_9934554_colorful-tree.html (visited on 10/07/2023).

Larreguy, Horacio and John Marshall. “The Effect of Education on Civic and Political Engagement in Nonconsolidated Democracies: Evidence from Nigeria”. In: *The Review of Economics and Statistics* 99.3 (2017), pp. 387–401. DOI: 10.1162/rest_a_00633.

Liang, Kung-Yee and Scott L. Zeger. “Longitudinal Data Analysis for Discrete and Continuous Outcomes”. In: *Biometrics* 42.1 (1986), p. 121. DOI: 10.2307/2531248.

Marshall, John. “Education and Voting Conservative: Evidence from a Major Schooling Reform in Great Britain”. In: *The Journal of Politics* 78.2 (2016), pp. 382–395. DOI: 10.1086/683848.

— “The Anti-Democrat Diploma: How High School Education Decreases Support for the Democratic Party”. In: *American Journal of Political Science* 63.1 (2019), pp. 67–83. DOI: 10.1111/ajps.12409.

Mayzlin, Dina, Yaniv Dover, and Judith Chevalier. “Promotional Reviews: An Empirical Investigation of Online Review Manipulation”. In: *American Economic Review* 104.8 (2014), pp. 2421–2455. DOI: 10.1257/aer.104.8.2421.

Meghir, Costas, Mårten Palme, and Emilia Simeonova. “Education and Mortality: Evidence from a Social Experiment”. In: *American Economic Journal: Applied Economics* 10.2 (2018), pp. 234–256. DOI: 10.1257/app.20150365.

Meyer, Andrew G. “The impact of education on political ideology: Evidence from European compulsory education reforms”. In: *Economics of Education Review* 56 (2017), pp. 9–23. DOI: 10.1016/j.econedurev.2016.11.003.

Miklikowska, Marta, Roderik Rekker, and Ales Kudrnac. “A Little More Conversation A Little Less Prejudice: The Role of Classroom Political Discussions for Youth’s Attitudes toward Immigrants”. In: *Political Communication* (2022), pp. 1–23. DOI: 10.1080/10584609.2022.2032502.

Milbrath, Lester and M. L. Goel. *Political participation : how and why do people get involved in politics*. Chicago: Rand McNally College Pub. Co, 1977. ISBN: 0528650955.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. "Does education improve citizenship? Evidence from the United States and the United Kingdom". In: *Journal of Public Economics* 88.9-10 (2004), pp. 1667–1695. DOI: 10.1016/j.jpubeco.2003.10.005.

Mincer, Jacob. *Schooling, Experience, and Earnings*. Natonal Bureau of Economic Research, 1974. ISBN: 0-870-14265-8.

Nie, Norman H., Jane Junn, and Kenneth Stehlik-Barry. *Education and Democratic Citizenship in America*. University of Chicago Press, 1996. 290 pp. ISBN: 0226583899.

Niederle, Muriel. "The Handbook of Experimental Economics, Volume Two". In: ed. by John H. Kagel and Alvin E. Roth. Princeton University Press, Jan. 2016. Chap. 8, pp. 481–562.

Niederle, Muriel and Lise Vesterlund. "Do women shy away from competition? Do men compete too much?" In: *The Quarterly Journal of Economics* 122.3 (2007), pp. 1067–1101.

— "Gender and Competition". In: *The Annual Review of Economics* 3 (2011), pp. 601–630.

Nyarko, Yaw and Andrew Schotter. "An Experimental Study of Belief Learning Using Elicited Beliefs". In: *Econometrica* 70.3 (2002), pp. 971–1005. DOI: 10.1111/1468-0262.00316.

Ottaviani, Marco and Peter Norman Sørensen. "Reputational Cheap Talk". In: *The RAND Journal of Economics* 37.1 (2006), pp. 155–175.

Pavlen. *Head with gears image. Intelligenz – Stockfoto*. ID: 109153217. Jan. 11, 2009. URL: https://www.istockphoto.com/de/foto/intelligenz-gm109153217-8183124?irgwc=1&cid=IS&utm_medium=affiliate&utm_source=photostockeditor&clickid=Tax3hiyHpxyLT-ewUx0Mo3EtUkBSap1tgySjT80&utm_content=258824&irpid=1327686 (visited on 10/07/2023).

Persson, Mikael. "Testing the Relationship Between Education and Political Participation Using the 1970 British Cohort Study". In: *Political Behavior* 36.4 (2014), pp. 877–897. DOI: 10.1007/s11109-013-9254-0.

Persson, Mikael, Karl-Oskar Lindgren, and Sven Oskarsson. "How does education affect adolescents' political development?" In: *Economics of Education Review* 53 (2016), pp. 182–193. DOI: 10.1016/j.econedurev.2016.03.015.

Phillips, Lawrence D. and Ward Edwards. "Conservatism in a Simple Probability Inference Task". In: *Journal of Experimental Psychology* 72.3 (1966), pp. 346–354.

Raven, J. C., J. Raven, and J. H. Court. *A manual for Raven's Progressive Matrices and Vocabulary Scales.* London: H. K. Lewis, 1998.

Rey-Biel, Pedro. "Equilibrium play and best response to (stated) beliefs in normal form games". In: *Games and Economic Behavior* 65.2 (2009), pp. 572–585. DOI: 10.1016/j.geb.2008.03.003.

Rockwell, Richard C. "Historical Trends and Variations in Educational Homogamy". In: *Journal of Marriage and the Family* 38.1 (1976), p. 83. DOI: 10.2307/350552.

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb. "Fast and wild: Bootstrap inference in Stata using boottest". In: *The Stata Journal: Promoting communications on statistics and Stata* 19.1 (2019), pp. 4–60. DOI: 10.1177/1536867x19830877.

Schlag, Karl and James Tremewan. "Simple belief elicitation: An experimental evaluation". In: *Journal of Risk and Uncertainty* 62.2 (2021), pp. 137–155. DOI: 10.1007/s11166-021-09349-6.

Schneider, Ulrich C. "Identifying and Estimating Beliefs from Choice Data - An Application to Female Labor Supply". In: *Working paper* (Feb. 2019).

Shurchkov, Olga. "Under pressure: Gender differences in output quality and quantity under competition and time constraints". In: *Journal of the European Economic Association* 10.5 (2012), pp. 1189–1213.

Smets, Kaat and Carolien van Ham. "The embarrassment of riches? A meta-analysis of individual-level research on voter turnout". In: *Electoral Studies* 32.2 (2013), pp. 344–359. DOI: 10.1016/j.electstud.2012.12.006.

Sobel, Joel. "A Theory of Credibility". In: *The Review of Economic Studies* 52.4 (1985), p. 557. DOI: 10.2307/2297732.

Solis, Alex. "Does Higher Education Cause Political Participation? Evidence from a Regression Discontinuity Design". In: *Working paper* (Aug. 2017).

Sondheimer, Rachel Milstein and Donald P. Green. "Using Experiments to Estimate the Effects of Education on Voter Turnout". In: *American Journal of Political Science* 54.1 (2010), pp. 174–189. ISSN: 0092-5853.

Statistik Austria. *Bildungsstand der Bevölkerung im Alter von 25 bis 64 Jahren, 1971 bis 2017*. Retrieved on July 30, 2020. June 2019. URL: https://www.statistik.at/wcm/idc/idcplg?IdcService=GET_NATIVE_FILE&RevisionSelectionMethod=LatestReleased&dDocName=020912.

Steele, C. M. and J. Aronson. "Stereotype threat and the intellectual test performance of African Americans". In: *Journal of personality and social psychology* 69.5 (1995), pp. 797–811.

Stelzeneder, Dominik. *Does Schooling Affect Political Attitudes? Quasi-Experimental Evidence*. Vienna Economics Papers. University of Vienna, Department of Economics, 2023. URL: https://EconPapers.repec.org/RePEc:vie:viennp:vie2301.

van Veldhuizen, Roel. “Gender Differences in Tournament Choices: Risk Preferences, Overconfidence or Competitiveness?” In: *Journal of the European Economic Association* (2022). DOI: 10.1093/jeea/jvac031.

Verba, Sidney, Nancy Burns, and Kay Lehman Schlozman. “Knowing and Caring about Politics: Gender and Political Engagement”. In: *The Journal of Politics* 59.4 (1997), pp. 1051–1072. DOI: 10.2307/2998592.

Weakliem, D. L. “The Effects of Education on Political Opinions: An International Study”. In: *International Journal of Public Opinion Research* 14.2 (2002), pp. 141–157. DOI: 10.1093/ijpor/14.2.141.

Wieland, Alice and Rakesh Sarin. “Domain specificity of sex differences in competition”. In: *Journal of Economic Behavior and Organization* 83 (2012), pp. 151–157.

WKO. *Lehrlingsstatistik - Demographische Entwicklung (Anteil Lehranfänger): 1970-2019*. Retrieved on July 30, 2020. Wirtschaftskammer Österreich. 2020. URL: http://wko.at/statistik/jahrbuch/LL_Demo.xlsx?_ga=2.84648094.642197588.1596111293-1677362337.1594654225.

Wolfinger, Raymond E. and Steven J. Rosenstone. *Who Votes?* Yale University Press, 1980.

Abstract

In the first chapter of my dissertation I study the direct causal effects of schooling on political attitudes of vocational students in Austria. I exploit that classes of apprentices of the same grade level and vocation are as good as randomly assigned to different school terms. This allows to compare apprentices who were at school for ten weeks with apprentices who were at work in their training firms during that time. I find that schooling has a positive direct causal effect on political interest of vocational students. This increase in political interest is, however, not accompanied by a significant increase in voting intention. Furthermore, my results suggest that apprentices who went to school while being exposed to a political affair support different parties than those apprentices who were exposed to the affair at work.

The second chapter is joint work with my co-authors Kai Barron and Sebastian Schweighofer-Kodritsch. Previous studies show that men are more competitive than women in stereotypically male tasks but not in stereotypically female tasks. However, to study the effect of a task's domain, these studies change the entire task and not the domain only. In our experiment we change the perceived domain of the task without changing the task itself: Participants solved *Raven's Progressive Matrices* that were either framed as *intelligence* tasks or as *creativity* tasks but strictly identical otherwise. This treatment affected the perceived gender stereotype of the task and both men's and women's competitive performance. The gender difference in competitiveness was, however, not affected by our treatment.

The third chapter is joint work with my co-author Boris Knapp. Misinformation in consumer markets - like fake reviews or biased endorsements - poses a challenge for consumers that can be conceptualized as a two-step problem. First, they need to form a belief about the probability that the information is biased. Second, they need to update their beliefs conditional on receiving information. This paper reports on an experiment that disentangles both steps utilizing a novel approach which relies neither on the estimation of individual utility parameters nor on belief elicitation. We find that inability to correctly compute posterior probabilities significantly impairs decisions which leads to a reduction in consumer surplus and that this effect is heterogeneous across gender.

Zusammenfassung

Im ersten Kapitel meiner Dissertation untersuche ich die direkten kausalen Effekte von Schulbesuch auf politische Einstellungen von Berufsschüler*innen in Österreich. Dabei mache ich mir zu Nutze, dass Berufsschüler*innen des selben Jahrgangs und Lehrberufs so gut wie zufällig in Klassen eingeteilt werden, die in unterschiedlichen Monaten unterrichtet werden. Dies erlaubt es, Berufsschüler*innen, welche zehn Wochen lang die Schule besucht haben, mit Berufsschüler*innen zu vergleichen, die während der gleichen Zeit im Lehrbetrieb gearbeitet haben. Meine Ergebnisse zeigen, dass Schulbesuch einen positiven direkten kausalen Effekt auf das politische Interesse von Berufsschüler*innen hat. Diese Zunahme an politischem Interesse ist jedoch nicht begleitet von einer signifikant höheren Intention, sich an einer Nationalratswahl zu beteiligen. Des Weiteren deuten meine Ergebnisse darauf hin, dass Berufsschüler*innen, welche die Berufsschule während der Enthüllung eines politischen Skandals besuchten, andere Parteien unterstützen als Berufsschüler*innen, die zur gleichen Zeit im Lehrbetrieb arbeiteten.

Das zweite Kapitel meiner Dissertation ist eine gemeinsame Arbeit mit meinen Co-Autoren Kai Barron und Sebastian Schweighofer-Kodritsch. Vorhergehende Studien zeigen, dass Männer kompetitiver als Frauen in stereotyp männlichen Aufgaben sind, aber nicht in stereotyp weiblichen Aufgaben. Um den Effekt der Domäne einer Aufgabe zu untersuchen verändern jene Studien jedoch die gesamte Aufgabe und nicht nur die Domäne der Aufgabe. In unserem Experiment verändern wir lediglich die wahrgenommene Domäne der Aufgabe, ohne die Aufgabe an sich zu verändern: Die Teilnehmer*innen unseres Experiments mussten *Raven's Progressive Matrices* lösen, welche entweder als *Intelligenz-* oder als *Kreativitätsaufgabe* geframed waren und sich sonst in nichts voneinander unterschieden. Dieses *treatment* beeinflusste das wahrgenommene Geschlechtsstereotyp der Aufgabe und die kompetitive Leistung sowohl der Männer als auch der Frauen. Der Geschlechtsunterschied bezüglich Kompetitivität wurde durch unser *treatment* jedoch nicht beeinflusst.

Das dritte Kapitel ist eine gemeinsame Arbeit mit meinem Co-Autor Boris Knapp. Falschinformationen in Verbrauchermärkten - wie beispielsweise falsche Produktbewertungen - stellen für Konsument*innen eine Herausforderung dar, welche als zweistufiges

Problem konzeptualisiert werden kann. Zuerst müssen die Konsument*innen abschätzen, wie hoch die Wahrscheinlichkeit ist, dass eine Information falsch ist. Zweitens müssen sie die bedingte Wahrscheinlichkeit berechnen, unter der Voraussetzung, dass sie eine konkrete Information bekommen. Für dieses Kapitel führten wir ein Experiment durch, anhand dessen wir beide Probleme entflechten können. Eine neuartige Herangehensweise erlaubt es uns, Aussagen über Fehlentscheidungen zu treffen, ohne individuelle Nutzenparameter zu schätzen oder subjektive Wahrscheinlichkeitseinschätzungen abzufragen. Unsere Ergebnisse zeigen, dass Fehler bei der Berechnung von bedingten Wahrscheinlichkeiten die Entscheidungen unserer Proband*innen signifikant verschlechtern. Dies führt zu einer signifikanten Reduktion ihrer Konsument*innenrente, wobei Männer und Frauen in unterschiedlichem Ausmaß betroffen sind.