Electronic supplementary material for

"On the social nature of eyes:

The effect of social cues in interaction and individual choice tasks."

Aurélien Baillon, Asli Selim, Dennie van Dolder

Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands

Table of Contents

I. Materials and methods 2
Recruitment of subjects 2
Website 2
Pictures 4
Structure of the website 6
Incentives 6
Questionnaire and additional data 6
II. Descriptive statistics of the subjects 8
III. Additional analyses - Joy of Destruction mini-game 9
IV. Additional analyses - Dictator game 12
V. Additional analyses - Ellsberg tasks 15
VI. Additional analyses - Simple vs. compound lotteries 21
References 24
Appendix A: Recruitment Emails 25
A.1. Recruitment e-mail 25
A.2. Reminder e-mail 26
Appendix B: Experimental Instructions 27
B.1. Welcome page 27
B.2. JoD mini-game 28
B.3. Dictator Game 29
B.4. Ellsberg Tasks 30
B.5. Compound vs. Simple Lotteries 31
B.6. Confirmation screen and additional questions 32
B.7. Final screen 34

I. Materials and methods

Recruitment of subjects

The experiment took place in the first half of June 2010. We recruited subjects by sending emails (in Dutch and English) to 400 students who were following a second-year Bachelor course called 'Research Project' at the Erasmus School of Economics (henceforth ESE) and to 200 other economics students (from the ESE) who indicated they would like to participate in experiments. Each student received a personalized link to the website developed for the experiment. Students were told that they had two weeks to participate if they wished and could receive up to €50. They received a reminder one week later. A copy of the English version of the invitation and the reminder email can be found in Appendix A. A total of 165 students completed the experiment. In section II we report descriptive statistics of the subjects.

Website

We constructed a website specifically for this experiment. The website was a replica of the ESE website. After the initial login to any computer at the ESE, Internet Explorer opens up automatically. The homepage consists of the ESE website, which displays news and important information. Students and staff members are required to use this website to look up information and for many administrative procedures. We used pieces of the HTML, CSS, and JavaScript codes of the ESE website to build our experimental website so as to copy its layout and functioning as closely as possible. In addition, we used PHP and SQL to generate the website on demand (subjects saw different versions of the website, since we randomized the conditions and the ordering of the tasks) and to record the answers of the subjects.

Similarly to the ESE website, our experimental website was bilingual (Dutch and English) and compatible with most browsers (such as Internet Explorer, Mozilla Firefox, Opera, Safari, and Chrome) and most screen sizes. One can see in Figure S1 and S2 that we slightly adapted the layout and removed some functionalities and links for the sake of clarity and readability. The website allowed subjects to stop the experiment at any time, and then go back and start exactly where they left off. Many subjects made use of this option.

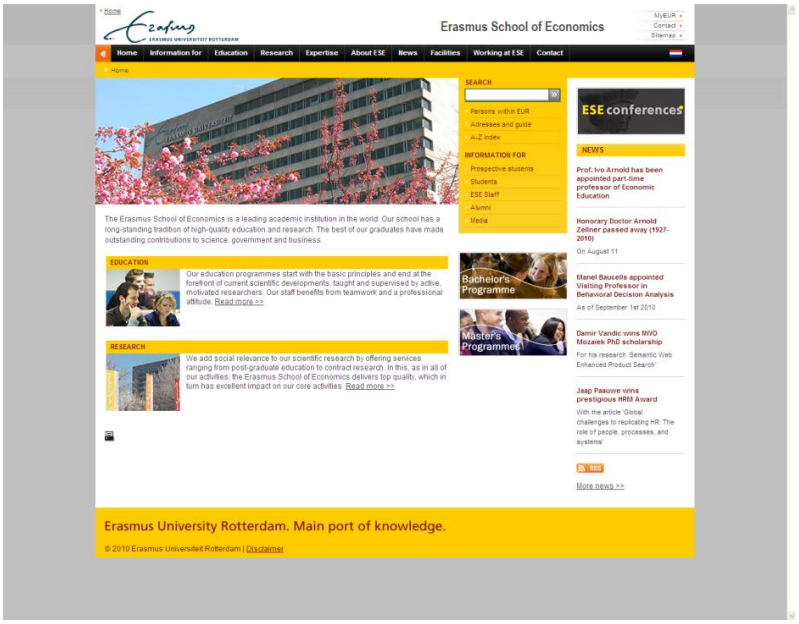

Figure S1: The original ESE website as seen with Internet Explorer in full screen mode.

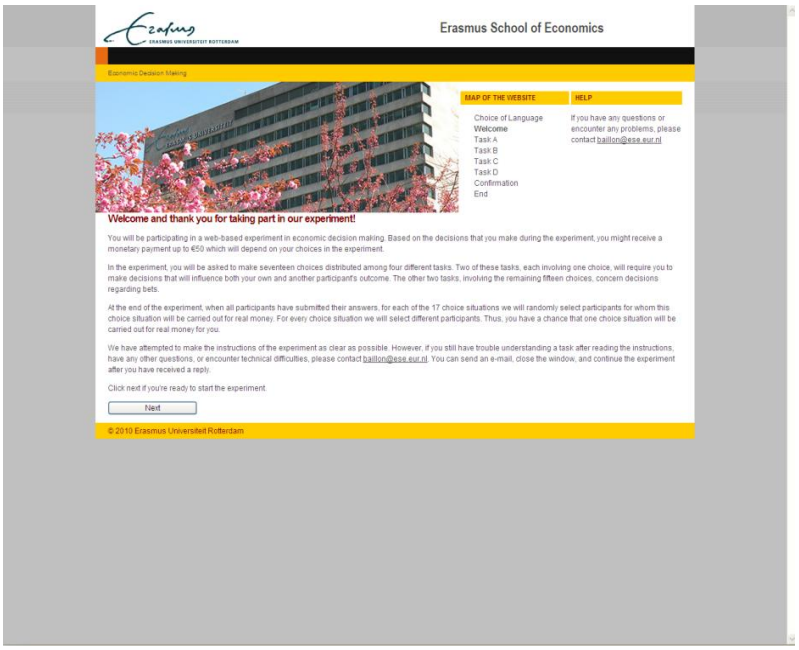

Figure S2: The welcome page of our experimental website as seen with Internet Explorer in full screen mode.

Pictures

To present our subjects with pictures of eyes and peers in an unobtrusive manner, we made use of the fact that the official ESE website has a banner displaying rotating pictures from the campus. Our own experience taught us that we do not really notice these pictures, unless we recognize somebody in them. Therefore, we took care that our subjects would not be likely to recognize anybody in the pictures employed in the experiment. Pictures on the ESE website are updated every three to six months and are adapted to the season or to special events. All ESE-related websites typically use the same design template, but adapt it by using their own, more specific, pictures.

In our study, the pictures displayed on the experimental website were the only differences between the three conditions (the eyes, the peers, and the control condition). We used four types of pictures, all of them being typical pictures one could find on the ESE website or any other university website:

- Common pictures (Figure S3 a and b): Two pictures depicting university buildings were taken from the official ESE website of Spring and Summer 2010. These pictures were mixed in with the condition-specific pictures, and thus common to all conditions. All the other pictures were taken by one of the authors for the purpose of the experiment.

- Control condition (Figure S3 c-g): We took photographs of empty classrooms and halls at the ESE to avoid reminders of peers and eyes of any kind.

- Peers condition (Figure S3 h-l): These pictures depict groups of students working, eating, walking, or talking to each other. We did not take these pictures at the ESE, but at other Dutch universities to minimize the risk that subjects might recognize somebody in the picture. We thought that recognition of oneself or of a friend in the pictures could draw the attention of a subject to the pictures and could possibly lead to behavioral effects beyond the scope of our current research. For the photographs that we took in other Dutch universities, we chose students and places such that they could have been from the ESE. Furthermore, we made sure that students in the pictures were not looking at the camera, to avoid a potential eyes effect. Please note that, in Figure S3, faces have been obscured for publication purposes but were visible in the experiment.

- Eyes condition (Figure S3 m-q): Previous studies on the effect of eyes employed a range of different stimuli, ranging from photographs of human eyes to highly stylized representations. In our study, we used photographs of the faces of statues of the school's

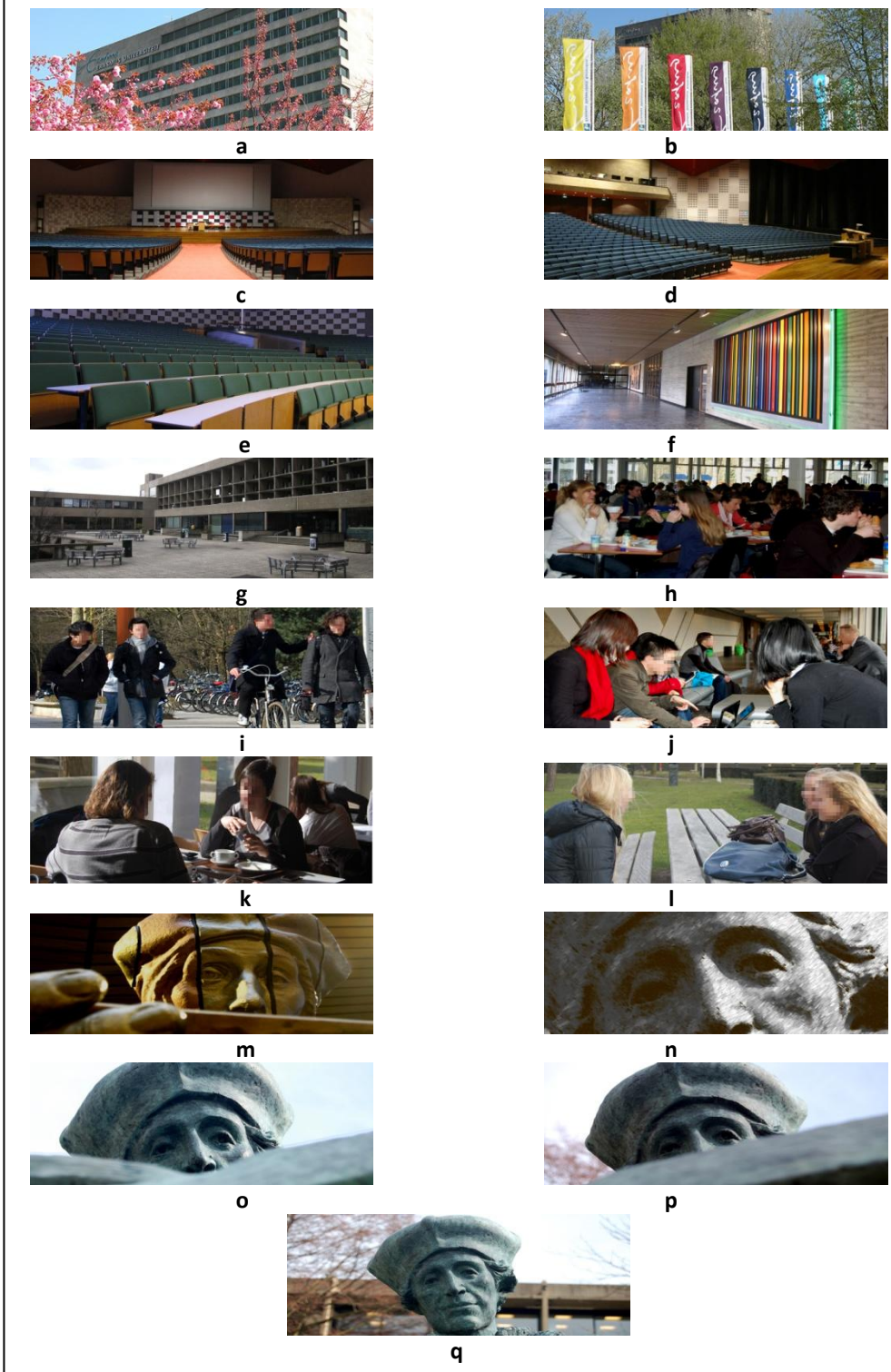

Figure S3. Pictures for the experimental website. Pictures a and b where common to all conditions. Pictures c through g were used in the control condition. Pictures h through l were used in the peers condition. Finally, pictures m through q were used in the eyes condition. In the experiment, faces on peers pictures were visible. They have been obscured here for publication purposes only.

namesake, Erasmus. The students are familiar with images of Erasmus because there are multiple statues of him on the campus and his image appears on official university documents. Thus, using such pictures would not appear out of the ordinary, and we could safely assume that the cues remained sufficiently subtle. Moreover, the neutral facial expressions displayed by the statues reduced the risk of accidently priming emotions

At the end of the experiment, subjects could make any comments that they wished. None of the comments mentioned the pictures. Each subject was randomly assigned to one of the conditions. If a subject left the website and then went back, the same picture set was displayed.

Structure of the website

On the first page of the experimental website, the subjects could choose whether the instructions would be displayed in Dutch or English. The second page (Appendix B.1.) briefly introduced the experiment and the payment mechanism (see below). The next four pages were dedicated to one of the four tasks employed, the order of which was randomized between subjects. These four pages were titled "Task A", "Task B", "Task C", and "Task D" independently of which task was described. We refer the reader to the main paper for the description of the tasks, and Appendix B.2.-B.5. for the English versions of the instructions used. At the end of the experiment, students were asked to fill in a brief questionnaire (see below and Appendix B.6.).

Incentives

In the experiment, we implemented a between-subject version of the Random Incentives System. We made clear in the experiment that each choice situation would be played out for real for some randomly selected subjects. The JoD mini-game and the dictator game involved two players each. There were 15 choice-situations for the Ellsberg task and the simple vs. compound lotteries task combined. We, therefore, randomly selected 19 subjects for whom the choices they made during the experiment would be carried out for real. After the experiment ended, these subjects were invited by e-mail to play out their choices for real and to receive the corresponding monetary awards.

Questionnaire and additional data

After completing the four decision tasks, were asked to fill in a short questionnaire about fear of negative evaluation (Leary, 1983). We have decided to disregard this questionnaire for two reasons.

First, subjects complained about it in their comments after the experiment (whereas most of the other comments were positive). The main problems seemed to be that these were the only psychological questions we used, which made it overtly obvious to the subjects what we were trying to measure and that the Dutch version of the scale was completely unidirectional (none of the questions were reversely coded, a higher score always implied more fear). Therefore, subjects considered the questions to be suggestive and disliked providing answers. Second, and possibly related to the first point, we noticed in the website's log files that many subjects preferred not to answer these questions and only did so when they were asked to do it for a second time. Therefore, it is very likely that they did not put much effort into answering the questions.

We also asked subjects whether they used a calculator during the experiment and a few demographic questions (age, gender, year of study, and nationality). Finally subjects could make any comments they wished.

Using internet technologies allowed us to gather extra data: IP addresses and the time subjects spent on each task. There is no unidirectional interpretation of these variables, however, so we omit them from our analyses. For instance, an IP address of the ESE could mean that the respondent was on the campus or was connected to the VPN of the ESE. Similarly, spending five minutes on one of the tasks could indicate that the subject gave it a lot of thought or simply that she was busy doing something else (like chatting or sending email) in the meantime and thus signal that the subjects was not thinking hard about the task at all.

Table S1: Descriptive statistics of the subjects

The table shows descriptive statistics for our sample of 162 subjects who participated in the experiment and answered the questionnaire (three subjects neglected to do so). *Age* is the subject's age measured in years. *Gender, Nationality, Calculator*, *Bachelor 1*, *Bachelor 2*, *Bachelor 3*, *Master*, and *Other* are dummy variables taking the value 1 if a contestant is female (*Gender*), Dutch (*Nationality*), indicates having used a calculator (*Calculator*), is a first year Bachelor student (*Bachelor 1*), a second year Bachelor student (*Bachelor 2*), a third year Bachelor student (*Bachelor 3*), a master student (*Master*), or indicates that she is neither in the first three years of her Bachelor, nor a Master student (*Other*), respectively.

	N	Mean	Median	St.Dev.	min	max
Age	162	21.10	21	2.06	18	33
Gender (Female = 1)	162	0.32	0	0.47	0	1
Nationality (Dutch = 1)	162	0.65	1	0.48	0	1
Calculator (yes = 1)	162	0.45	0	0.50	0	1
Year of study						
Bachelor 1	162	0.14	0	0.35	0	1
Bachelor 2	162	0.57	1	0.50	0	1
Bachelor 3	162	0.12	0	0.32	0	1
Master	162	0.16	0	0.37	0	1
Other	162	0.01	0	0.11	0	1

II. Descriptive statistics of the subjects

Table S1 displays the descriptive statistics for the 162 subjects that completed the questionnaire at the end of the experiment (three subjects neglected to do so). The majority of our subjects were male (68%), Dutch (65%), and in the second year of their bachelor's degree (57%). Furthermore, around 14% were in their first year of the bachelor's degree, 12% were in their third year, 16% were following a master's program, and 1% did not fall into any of these categories. Both the average and the median age were 21, and age ranged from 18 to 33. It should, however, be mentioned that over 90% of our subjects was under the age of 25 (not in table). A considerable share of subjects admitted to having used a calculator during the experiment (45%). Note that using a calculator was by no means forbidden in the experiment. We simply asked this question since using a calculator would facilitate finding correct answers in one of the tasks.

Table S2: Probit analyses on destruction rate in the JoD mini-game

The table displays results from the Probit regression analyses of subjects' decisions to destroy (1) or not destroy (0) part of another subject's endowment. *Eyes* and *peers* are dummy variables taking the value 1 if subjects were in the eyes condition (*Eyes)* or the peers condition (*Peers*), respectively. Definitions of the other variables are as in Table S1. Model S2.1 is estimated on the entire sample of 153 subjects who successfully submitted a decision in the JoD mini-game. Model S2.2 and Model S2.3 are estimated on the set of 150 subjects who both successfully submitted a decision in the JoD mini-game and answered the questionnaire at the end of the experiment. Model S2.4 is estimated on the set of 104 subjects for whom the decision in the JoD mini-game was successfully recorded the first time round (i.e. not affected by technical problems) and who answered the questionnaire at the end of the experiment. For each explanatory variable, the marginal effect at the covariate means is shown. Robust standard errors are used and *p*-values are shown in parentheses.

	Probability destroying							
	Model S2.1		Model S2.2		Model S2.3		Model S2.4	
Condition dummies (Control is reference)								
Eyes	-0.18	(0.01)	-0.18	(0.01)	-0.18	(0.01)	-0.18	(0.03)
Peers	-0.17	(0.02)	-0.14	(0.04)	-0.15	(0.03)	-0.15	(0.08)
Control variables								
Age			-0.01	(0.50)	-0.01	(0.55)	0.00	(0.90)
Gender (female = 1)			0.06	(0.46)	0.07	(0.40)	0.05	(0.64)
Nationality (Dutch = 1)			-0.26	(0.00)	-0.26	(0.00)	-0.30	(0.00)
Year of study (First year and other are reference)								
Bachelor 2			0.02	(0.86)	-0.04	(0.76)	-0.23	(0.24)
Bachelor 3			0.04	(0.80)	0.01	(0.93)	-0.21	(0.00)
Master			0.18	(0.39)	0.14	(0.49)	-0.04	(0.87)
Calculator (yes = 1)			-0.06	(0.43)	-0.06	(0.42)	0.01	(0.93)
Second time (yes = 1)					-0.08	(0.27)		
LL		-82.15		-69.21		-68.74		-47.2
N		153		150		150		104

III. Additional analyses - Joy of Destruction mini-game

In the JoD mini-game, our subjects had the option to pay €1 to destroy €10 of another player's endowment. As discussed in the main paper, χ^2-tests showed that subjects are significantly less likely to destroy the endowment of the other subject both in eyes condition and the peers condition, relative to the control. There is no difference between the eyes and the peers condition. In the current section we show that these results are robust when we apply Probit regressions and control for the effect of other variables. Furthermore, due to a technical problem, the decisions submitted by some of the subjects (58 out 165) were initially not recorded in the database. These subjects received an email telling them that they could go back to the website to fill in the missing decision and most of them (46 out of 58) did so. Here, we provide additional tests showing that there is no indication that this data problem influenced our results.

Table S3: Descriptive statistics on destruction rate

The table shows descriptive statistics on destruction rate in the JoD mini-game depending on whether subjects' decisions were recorded the first time the subjects submitted them, or whether subjects had to record their decisions for a second time due to a technical problem with the website. Results are shown for all conditions both separately and combined. Overall statistics are provided in the final column.

	First time		Second time		Overall	
	N	% destroy	N	% destroy	N	% destroy
Eyes	33	18.18	18	16.67	51	17.65
Peers	41	21.95	12	8.33	53	18.87
Control	33	39.39	16	37.50	49	38.78
Total	107	26.17	46	21.74	153	24.84

Table S2 shows the results of a Probit model on the probability that a subject destroys the endowment of another subject; significance levels are based on robust standard errors. Furthermore, since coefficients in a Probit model do not offer intuitive interpretations in terms of effect size, we report marginal effects evaluated at the covariate means. Model S2.1 provides a simple comparison between conditions. In line with the χ^2-tests, we observe that destruction rates both in the eyes (P = 0.01) and the peers (P = 0.02) conditions are significantly lower than in the control condition. There is no significant difference between the eyes and the peers condition (P = 0.87, untabulated). Adding our control variables (Table S2, Model S2.2), we find that only nationality has a significant influence on destruction. Dutch students are significantly less likely to destroy the other's endowment (P < 0.01). The effect of both the eyes and the peers condition remain statistically significant (respectively P = 0.01 and P = 0.04). In short, these analyses allow us to conclude that the simple, non-parametric tests applied in the main article prove robust in more advanced analyses controlling for age, gender, nationality, education year, and the use of a calculator.

Table S3 shows descriptive statistics on destruction rates depending on whether the subject's decision was recorded the first time round or whether they had to record their decisions a second time due to the data storage problem. For the eyes and the control condition, the findings are highly similar in both cases. Subjects destroy the others' endowment in 18.18% of the cases when recording their decision for the first time and 16.67% when recording it the second time in the eyes condition. These statistics are 39.39% and 37.50%, respectively, for the control condition. When investigating the peers condition the gap appears a bit larger, subjects destroy the other's endowment in 21.95% of the cases when answering the question for the first time, and 8.33% of the

cases when answering the question for a second time. A Fisher's exact test, however, indicates that this difference is not statistically significant ($P = 0.42$).

Overall, the qualitative pattern seems to be the same independent of whether subjects recorded their decisions for the first or the second time: subjects in the eyes and peers conditions destroy at a similar rate, which is lower than the destruction rate in the control condition. Performing χ^2-tests, we find that even for the subset of subjects who recorded the questions for the first time the differences between conditions approach significance (comparing eyes with control: $P = 0.06$; comparing peers with control: $P = 0.10$). The additional observations from the subjects who had to record their decision for a second time thus only strengthen the statistical evidence for an already apparent pattern.

Table S2, Model S2.3 incorporates a dummy variable taking the value 1 if the decision had to be recorded a second time, 0 otherwise, into the full model. This analysis shows all the results to be robust and that being requested to answer the question a second time does not steer behavior in a particular direction. Model S2.4 reports the results of estimating the model on the subset of subjects whose decisions were successfully stored the first time round. While the significance levels drop a bit, we observe that the estimates of the marginal effects are not at all affected by leaving out the subjects who had to record their decisions for a second time. This provides further indication that the data storage problem did not affect our results in a meaningful way.

Table S4: Regression results on giving in the dictator game

The table displays results from regressions on the subjects' giving behavior in the dictator game. Model S4.1 and Model S4.2 display results of Tobit regression analyses on the amount donated by the subjects. Model S4.3 and Model S4.4 displays results of Probit analyses on the subjects' decisions to either donate (1) or not (0). Definitions of the variables are as in previous tables. In the results of the Probit regressions we depict marginal effects at the covariate means. For both Tobit and Probit models, we apply robust standard errors. *P*-values are shown in parentheses.

	Amount transferred				Probability giving			
	Model S4.1		Model S4.2		Model S4.3		Model S4.4	
Constant	5.15	(0.06)	-3.16	(0.91)				
Condition dummies (Control is reference)								
Eyes	6.31	(0.08)	7.41	(0.04)	0.13	(0.13)	0.18	(0.05)
Peers	-2.91	(0.48)	-2.81	(0.51)	-0.12	(0.18)	-0.09	(0.38)
Control variables								
Age			0.88	(0.50)			0.01	(0.72)
Gender (female = 1)			-3.04	(0.37)			0.04	(0.64)
Nationality (Dutch = 1)			-5.62	(0.09)			-0.22	(0.01)
Year of study (First year and other are reference)								
Bachelor 2			-5.29	(0.29)			-0.10	(0.39)
Bachelor 3			-4.92	(0.39)			0.06	(0.73)
Master			-7.31	(0.35)			-0.17	(0.41)
Calculator (yes = 1)			-3.60	(0.31)			-0.02	(0.85)
Sigma	18.80		18.51					
LL	-490.89		-474.73		-104.24		-97.92	
N	165		162		165		162	

IV. Additional analyses - dictator game

As mentioned in the main article, the results from non-parametric Mann-Whitney tests suggest that the amount donated in the dictator game differs significantly between conditions. In particular, subjects donate significantly more to the other subject in the eyes condition compared to the other two conditions. There is no significant difference in the amount donated between the peers and the control condition. Looking at the percentage of subjects who decide to give away money, χ^2-tests indicate that neither the eyes nor the peers conditions differs significantly from the control, but that subjects in the eyes condition are significantly more likely to donate as compared to subjects in the peers condition. In the present section we will show that these results are robust, or even strengthened, by performing more advanced analyses and controlling for the effect of other variables on the willingness to donate money to a stranger.

First, we analyze the amount donated by the Dictator by means of a Tobit model. We use a Tobit model to account for the fact that our dependent variable "Amount given" is censored between €0 and €50. Model S4.1 and Model S4.2 show our results, significance levels being based on robust standard errors. Model S4.1 presents a simple test of the condition effects to compare the results from the (parametric) Tobit analyses with those of the (non-parametric) Mann-Whitney tests reported above. We observed that the difference between the eyes condition and the control condition decreases in significance due to the distributional assumptions made in the Tobit. Still, the difference between the eyes and the control conditions remains marginally significant (P = 0.08)—and is significant if we perform a one-sided test (P = 0.04). More importantly, however, adding our control variables (Model S4.2) the condition effect increases in significance, becoming significant at the 5% level in a two-sided test (P = 0.04). None of the control variables seems to have a strong influence on behavior, except that Dutch students seem less willing to donate money (P = 0.09). Interpreting the parameters, an individual's willingness to donate increases by about €7.41 euro's in the eyes condition compared to the control condition. This difference is larger than the observed difference in money allocated between conditions due to the fact that the Tobit takes censoring in the data into account. As could be expected, the difference between the eyes condition and the peers condition is significant both in Model S4.1 and Model S4.2 (P < 0.02, untabulated).

Model S4.3 and S4.4 show the results of Probit models on the probability that a subject allocates a non-zero amount to another subject. As before, we report marginal effects evaluated at covariate means and significance levels are based on robust standard errors. Model S4.3 provides a simple comparison between conditions. As suggested by the $\chi 2$-tests reported earlier, the Probit results shows that no condition differs significantly from the control condition, while the subjects in the eyes condition are significantly more likely to donate compared to those in the peers condition (P < 0.01, untabulated). Adding our control variables, however, increases the significance of the eyes condition sharply, indicating that subjects in the eyes condition are significantly more likely to give a positive amount to another subject compared to the subjects in the control condition (p = 0.05). The size of this effect is impressive: the subjects in the eyes condition are almost 18 percentage points more likely to donate money compared to the subjects in the control condition and more than 25 percentage points more likely to donate money compared to the subjects in the peers condition (P < 0.01). The difference between the peers condition and the control condition remains insignificant (P = 0.38). Again the only control variable that seems to matter is nationality, Dutch students are 22 percentage points less likely to allocate a positive amount to another subject (P = 0.01).

In conclusion, using more advanced analyses and controlling for a range of other variables that can potentially influence giving behavior, we find that this only strengthens the conclusions drawn in the main article. It is interesting to note that while Dutch students acted less anti-social in the JoD mini-game, these same students acted less pro-social in the Dictator game. This suggests that this subcategory of students is not, in fact, more or less kind, but rather is less likely to deviate from the prediction of rational self-interest.

Table S5: Probit analyses on choosing risk over ambiguity

The table displays results from the Probit regression analyses of subjects' decisions to choose the risky Bag K (1) over the ambiguous Bag U (0). Definitions of the variables are as in previous tables. For each explanatory variable, the marginal effect is shown at the covariate means. Robust standard errors are used and *p*-values are shown in parentheses.

	Probability Choosing Bag K			
	Model S5.1		Model S5.2	
Condition dummies (Control is reference)				
Eyes	-0.07	(0.46)	-0.11	(0.27)
Peers	-0.21	(0.02)	-0.19	(0.03)
Control variables				
Age			-0.03	(0.17)
Gender (female = 1)			0.11	(0.14)
Nationality (Dutch = 1)			0.06	(0.43)
Year of study (First year and other are reference)				
Bachelor 2			0.06	(0.58)
Bachelor 3			0.08	(0.49)
Master			0.03	(0.84)
Calculator (yes = 1)			0.05	(0.49)
LL	-85.79		-81.58	
N	165		162	

V. Additional analyses – Ellsberg tasks

Here we present a number of additional analyses regarding the ambiguity questions used in the experiment. First, we will show that the findings in the general Ellsberg tasks (Ellsberg, 1961) that are reported in the main article are robust if we use Probit analyses and control for the effect of other variables on the subjects' decisions. Second, we will investigate the other questions posed to subjects. As mentioned in the main article, we implemented the standard Ellsberg choice situation with a 50-50 proportion of red and black chips in Bag K, but we also varied the proportion of red and black chips from 10%-90% to 90%-10% (i.e., 10%-90%, 20%-80%, 30%-70%…). Here we will show that when the probability was different from 50%, subjects overwhelmingly select the normatively superior option, i.e., Bag K if the probability of winning in this bag is 60% or higher, Bag U if the probability of winning in Bag K is 40% or lower. As a result, no clear differences between conditions can be detected in these scenarios.

As discussed in the main article, the standard Ellsberg question we employed involved two bags containing black and red chips; in one bag (Bag K) the proportion of red and black chips was known, whereas in the second bag (Bag U) this proportion was not known. The subjects were asked to

choose a color and a bag to draw a chip from. If the color of the drawn chip matched the one they had chosen, they received €50. When the proportion of red and black chips is 50-50, Bag K and Bag U are normatively equivalent, but many studies have shown that a disproportionate number of people choose Bag K (Camerer & Weber, 1992). In line with this common pattern, we observe that 85.5% of the subjects chose Bag K in our control condition. Using χ^2-tests, we find that behavior does not differ between the eyes and the control condition. In the peers condition, however, subjects are significantly less likely to show a bias in favor of Bag K.

To investigate the robustness of this finding, we perform Probit analyses on the likelihood of choosing Bag K. The findings are reported in Table S5. As before, we report marginal effects around covariate means and apply robust standard errors in order to calculate statistical significance. These analyses yield results that are perfectly in line with the $\chi 2$-tests reported in the paper. That is, subjects are significantly less likely to show a bias toward bag K in the peers condition as opposed to the control condition ($P < 0.03$). The difference between eyes and peers turns marginally significant in Model S5.1 ($P = 0.09$, untabulated), but drops in significance when background characteristics are accounted for ($P = 0.32$, untabulated). The difference between the eyes and the control condition does not reach significance in any of the models ($P > 0.27$). None of the control variables influences the choice for Bag K. These analyses thus show the effect of the peers condition, as compared to the control condition, to be a rather robust phenomenon, whereas there is no evidence for an effect of the "eye" condition.

As mentioned above, we also asked subjects to choose between the ambiguous prospect and a range of risky prospects with a probability of winning of 10, 20, 30, 40, 60, 70, 80, and 90%. We start by using these questions to create an index of "ambiguity aversion", defined as the degree to which people tended to prefer the risky prospect to the ambiguous one. We generate this index by counting the number of times a subject prefers the risky prospect over all nine choice-tasks. The higher this index, the greater the degree to which the subject shows a preference for the risky prospect over the ambiguous one. This index indicates, as shown in Table S6, that subjects in the peers condition were less attracted by the risky urn. This difference is significant in Mann-Whitney tests (in comparison with the control: $z = 1.99$, $P = 0.05$, comparison with the eyes condition: $z = 1.96$, $P = 0.05$). The difference between the other two conditions is not significant ($z = 0.02$, $P = 0.98$).

Table S6: Descriptive statistics of the ambiguity aversion index over different conditions

The table displays the descriptive statistics of the ambiguity aversion index over different conditions. We calculated the index by counting the number of times a subject prefers the risky prospect over all nine choice-tasks. The higher this index, the greater the degree to which the subject shows a preference for the risky prospect over the ambiguous one.

	N	Mean	Median	St.Dev.	min	max
Eyes	55	5,11	5,00	1,29	0	9
Peers	55	4,75	5,00	1,27	0	9
Control	55	5,02	5,00	0,91	2	7
Total	165	4,96	5,00	1,17	0	9

An interesting point, however, is that the above effect seems to be caused entirely by the choice when the probability of winning in the risky prospect is 50%. When we leave out this choice in our construction of the index, we find no significant differences between groups ($P > 0.21$). Figure S4 illustrates this point: when it comes to the index, the major differences arise around a score of four or five. It should be noted that in this task most subjects (87.3%) act consistently; they stick to the ambiguous prospect until the probability of winning in the risky prospects becomes sufficiently high, and after this point they consistently choose the risky prospects and do not switch back to the ambiguous one. Therefore, the switching points at which subjects decides to give up the ambiguous prospect for the risky ones drive the difference in the indexes that we observe between the conditions. Switching at the 50% risky prospect implies a score of five, switching prior to it at the 40% prospect implies a score of six, and switching only at the 60% prospect implies a score of four. Therefore, as can be clearly seen in Figure S4, the choice at the 50% prospect is the main driving force behind the differences in the ambiguity index.

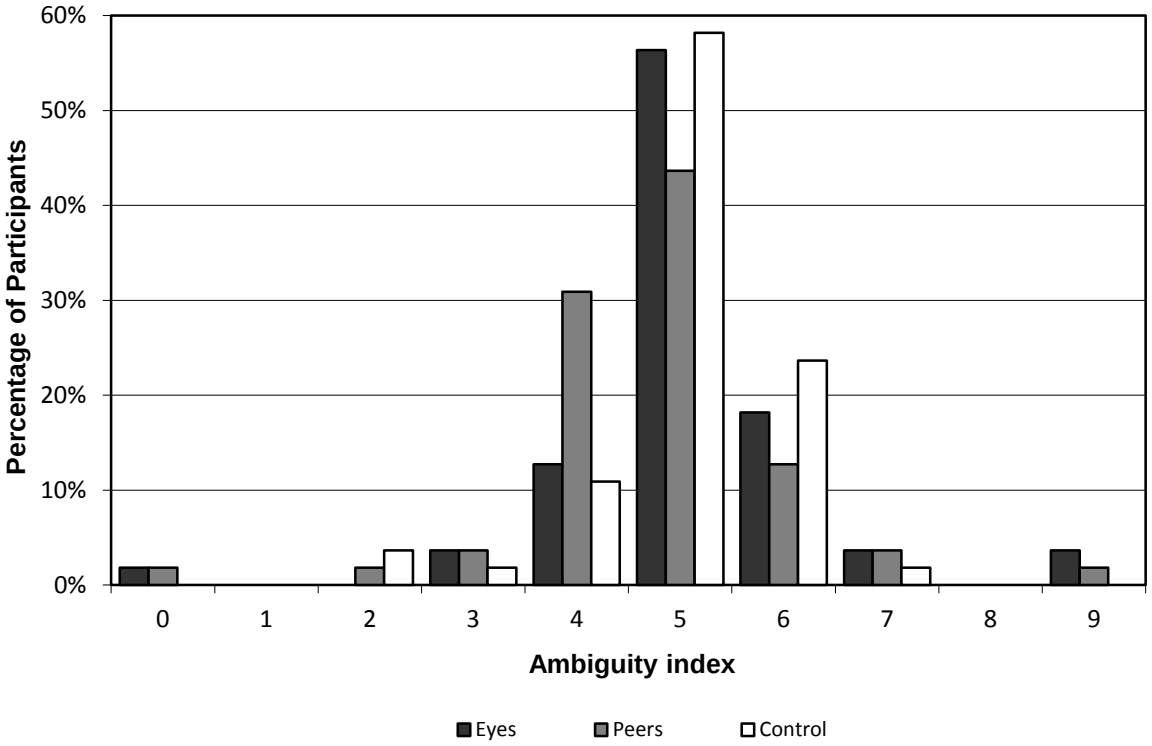

Figure S4: Values on the ambiguity aversion index in the three conditions. The ambiguity aversion index is calculated by counting the number of times a subject prefers the risky prospect over all nine choice-tasks. The higher this index, the greater the degree to which the subject shows a preference for the risky prospect over the ambiguous one.

Finally, Figure S5 shows this finding by depicting the percentage of subjects who chose a risky prospect as a function of the probability of winning in that risky prospect. It is easy to see that when the probability of winning in the risky prospect is not 50%, most of the subjects show a strong preference for either of the two prospects: when the probability is lower than 50%, a strong majority of subjects choose the ambiguous prospect, and when it is higher than 50%, an overwhelming majority of subjects choose the risky one. Due to these strong majorities, we can no longer use $\chi 2$-tests to statistically test for differences between conditions in these tasks, as the $\chi 2$-test is not reliable when data is highly unbalanced. Therefore, we employ Fischer's exact test to test for differences between conditions in the choice task. It should be noted that applying Fischer's exact test does not alter our conclusions for choice regarding the 50% prospect, the difference between the peers and the control condition remains highly significant (P = 0.03), although the difference between the peers and the eyes condition is no longer significant (P = 0.13). For the other choice-tasks, neither the eyes nor the peers condition differs significantly from the control (P > 0.11). The only condition comparison that approaches significance is that between eyes and peers in the 90% choice task (P = 0.05).

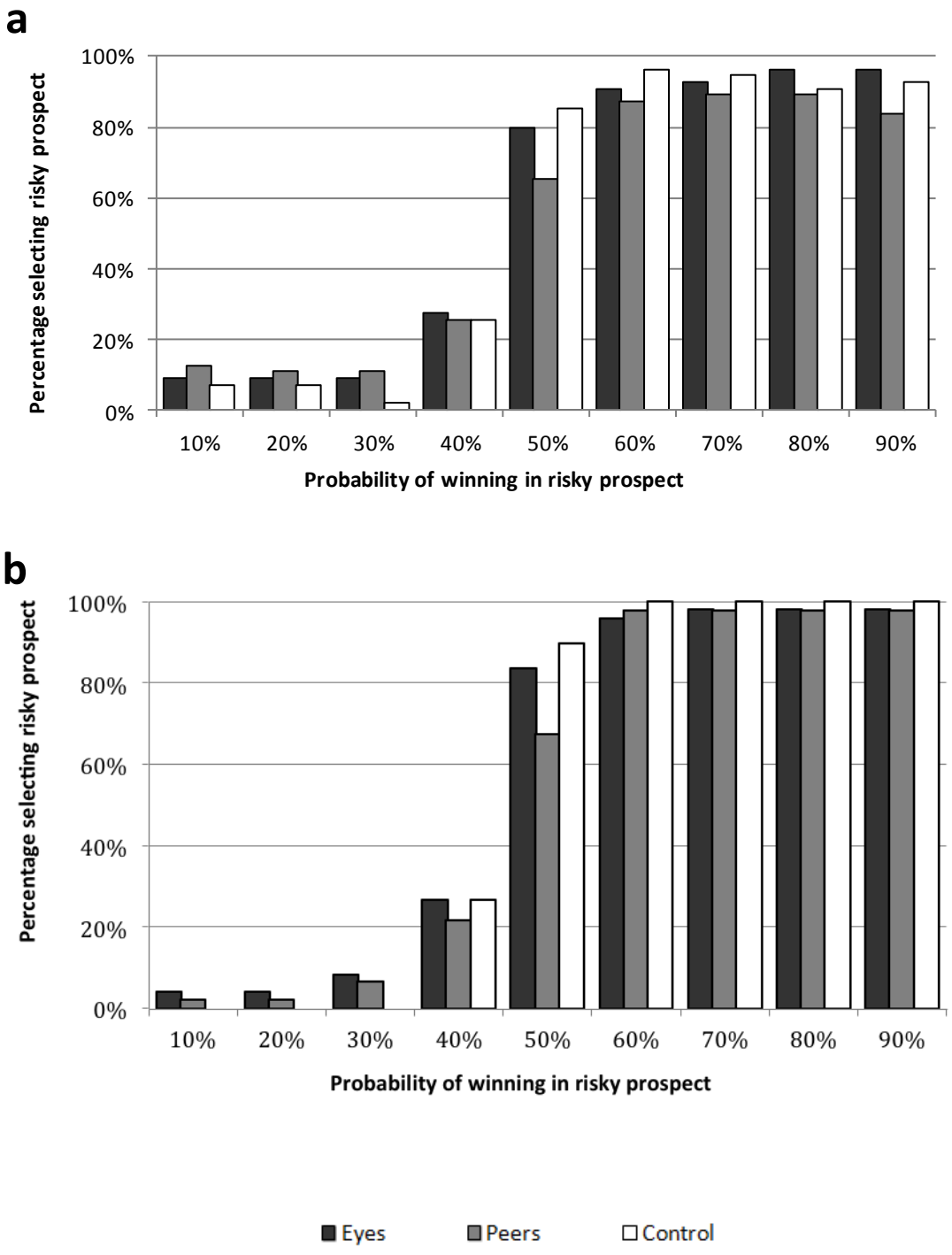

Figure S5: Likelihood of choosing the risky prospect as function of winning probability. The figures display the percentage of subject selecting the risky prospect for each of the nine choice questions that vary the probability of winning from 10% to 90% by condition. The Figure in (a) displays results for the complete set of data. The Figure in (b) displays results excluding a few subjects who showed inconsistent preferences in this task.

It should, however, be noted that this difference completely disappears when we only focus on those 87.3% of the subjects who behave completely consistent within this task. If we leave the inconsistent subjects out of the analyses, the eyes and peers condition yield the exact same propensity to choose the risky prospect in the 90% choice task (p = 1.00). Focusing on these consistent individuals, again only the difference between the peers and the control condition at the 50% choice becomes significant (P = 0.01), while the difference between the eyes and peers condition becomes marginally significant (P = 0.09). No further condition differences emerge (P > 0.11). This implies that the only robust pattern in the ambiguity aversion task is the finding that when choosing between an ambiguous prospect and a risky prospect with a 50% winning probability, subjects in the peers condition are significantly less likely to show a bias in favor of the risky prospect as compared to the subjects in the control condition.

Table S7: Probit analyses on the likelihood of making a mistake

The table displays results from the Probit regression analyses on the likelihood that subjects choose a compound gamble over a strictly better simple gamble at least once. Model S7.1 and Model S7.2 compare both the eyes and the peers condition to the control, Model S7.3 and Model S7.4 compare the peers condition to the two other conditions combined. Definitions of the variables are as in previous tables. For each explanatory variable, we report marginal effects evaluated at covariate means. Robust standard errors are used and *p*-values are shown in parentheses.

	Control as reference				Control + eyes as reference			
	Model S7.1		Model S7.2		Model S9.3		Model S9.4	
Condition dummies (Control is reference)								
Eyes	0.00	(1.00)	0.08	(0.42)				
Peers	-0.16	(0.08)	-0.23	(0.06)	-0.16	(0.04)	-0.27	(0.01)
Control variables								
Age			0.00	(0.87)			-0.01	(0.80)
Gender (female = 1)			0.19	(0.02)			0.19	(0.02)
Nationality (Dutch = 1)			-0.20	(0.02)			-0.19	(0.03)
Year of study (First year and other are reference)								
Bachelor 2			0.05	(0.71)			0.05	(0.72)
Bachelor 3			-0.03	(0.87)			-0.02	(0.90)
Master			0.15	(0.41)			0.17	(0.33)
Calculator (yes = 1)			-0.58	(0.00)			-0.58	(0.00)
LL		-107.66		-69.43		-107.66		-69.72
N		165		162		165		162

VI. Additional analyses - simple vs. compound lotteries

As shown in the main article, we find that there are significant differences in the likelihood that subjects mistakenly choose the compound gamble over the superior simple gamble. While there is no difference between the eyes and the control condition, χ^2-tests indicates that the likelihood of making such a mistake is marginally significantly lower in the peers condition as compared to the other two conditions separately, and significantly lower if we combine the other two conditions. Furthermore, looking at the number of mistakes reveals a similar pattern. Mann-Whitney tests indicate that subjects in the peers condition make marginally significantly fewer errors in the peers as opposed to the control condition, where the eyes condition does not differ significantly from the other two conditions. In the present section we will show that these results are robust, or even strengthened, when performing more advanced analyses and controlling for the effect of other variables on the likelihood of making errors.

Table S8: Ordinal Probit analyses on the number of mistakes

The table displays results from Ordinal Probit regression analyses on the number of mistakes that subjects make. The Condition model only includes condition dummies, whereas the Full model includes all our controls. We report both coefficients and marginal effects on the likelihood that a person makes a specific number of errors evaluated at covariate means. Robust standard errors are used and *p*-values are shown in parentheses.

	Coeff.	Sign.	Marginal Effects						
			P(0)	P(1)	P(2)	P(3)	P(4)	P(5)	P(6)
Condition									
Condition dummies (Control is reference)									
Eyes	-0.15	(0.48)	0.06	0.00	0.00	-0.01	-0.01	-0.02	-0.01
Peers	-0.43	(0.04)	0.16	0.00	-0.01	-0.03	-0.04	-0.05	-0.04
LL		-285.94							
N		165							
Full									
Condition dummies (Control is reference)									
Eyes	-0.10	(0.63)	0.04	0.00	0.00	-0.01	-0.01	-0.01	0.00
Peers	-0.52	(0.02)	0.20	0.01	-0.03	-0.06	-0.05	-0.04	-0.02
Control variables									
Age	-0.03	(0.66)	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Gender (female = 1)	0.16	(0.40)	-0.06	-0.01	0.01	0.02	0.02	0.01	0.01
Nationality (Dutch = 1)	-0.56	(0.01)	0.20	0.02	-0.02	-0.06	-0.06	-0.05	-0.03
Year of study (First year and other are reference)									
Bachelor 2	-0.01	(0.97)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bachelor 3	0.16	(0.64)	-0.06	-0.01	0.01	0.02	0.02	0.01	0.01
Master	0.14	(0.72)	-0.05	0.00	0.00	0.02	0.01	0.01	0.01
Calculator (yes = 1)	-1.51	(0.00)	0.53	0.02	-0.06	-0.15	-0.14	-0.11	-0.08
LL		-239.70							
N		162							

First, we perform a Probit analysis on the likelihood of making one or more mistakes. We consider four model: two in which we compare the eyes and peers condition to the control condition, and two in which we compare the peers condition to the other two. For both analyses we apply a simple model without control variables and a model that account for the effects of several control variables. Table S7 shows our results, again, using robust standard errors and reporting marginal effects evaluated at covariate means in order to give parameters a substantive meaning.

As Table S7 clearly shows the standard condition only models (Model S7.1 and Model S7.3), are perfectly in line with the $\chi 2$-tests reported earlier; we observe no difference between the eyes and the control, a marginal significant difference between peers and the other two conditions separate (P = 0.08) and a significant difference between the peers condition and the two other conditions combined (P = 0.04). Adding the control variables (Model S7.2 and Model S7.4), we find that these results are robust and indeed increase in significance somewhat (respectively P = 0.06 and P = 0.01).

Furthermore, females are more likely to make at least one error, Dutch student are less likely to do so, and the use of a calculator drastically decreases the likelihood of making an error.

Secondly, we estimate an Ordinal Probit model where the dependent variable is the number of mistakes (0 through 6). Table S8 shows our results. We report coefficients with their significance levels. As in all previous analyses, significance levels are based on robust standard errors. Furthermore, we report marginal effects evaluated at covariate means for each possible outcome category (0 through 6 mistakes). We present two models: a basic condition model without control variables, and a full model, which includes control variables alongside the general condition effects.

As can be seen in Table S8, both models show that the subjects in the peers condition are significantly less likely to make errors ($P < 0.04$). The difference between eyes and peers is insignificant in the first model ($P = 0.16$, untabulated) and marginally significant in the second model ($P = 0.06$, untabulated). With respect to the control variables, we find that Dutch students are significantly less likely to make mistakes, supporting the idea that this sub-group behaves more in line with rationality based arguments. Naturally, subjects who use calculators are also significantly less likely to make mistakes.

In conclusion, the non-parametric tests reported in the paper are in line with the more advanced analyses including control variables reported here. In general, adding control variables seems to strengthen our results rather than weaken them.

References

Camerer, C.F., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, 5, 325–370.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75, 643–669.

Leary, M.R. (1983). A brief version of the fear of negative evaluation scale. *Personality and Social Psychology Bulletin*, 9, 371-375.

Appendix A: Recruitment Emails

A.1. Recruitment e-mail

Dear student,

We would like to invite you to participate in a **web-based** experiment on economic decision-making, run by the "Behavioural Economics Group" at the ESE. The experiment is carried out online, so you can participate at any time and anywhere you like over the *next two weeks*. All you have to do is to use the link below and follow the instructions on the website. The experiment will take 10-15 minutes of your time, and in return you will get a chance to win up to 50 euros! We will randomly select 19 people among the participants and have a budget of 850 euros for this experiment.

Your personal link to the experiment is:

[PERSONALIZED LINK TO THE WEBSITE]

You will not have to log into our website: this personal link will automatically register that you have participated in the experiment.

The experiment will be online only 2 weeks, so if you want to have a chance of winning 50 euros, you should make sure to participate in the experiment very soon.

Thank you for your interest in our experiments!

Best regards,

Aurelien Baillon

A.2. Reminder e-mail

Dear student,

There is only one week left to take part in our web-based experiment and to get a chance to win 50 euros.

Your personal link to the experiment is:

[PERSONALIZED LINK TO THE WEBSITE]

More information:

This is a **web-based** experiment on economic decision-making, run by the "Behavioural Economics Group" at the ESE. The experiment is carried out online, so you can participate at any time and anywhere you like. All you have to do is to use the link above and follow the instructions on the website. The experiment will take 10-15 minutes of your time, and in return you will get a chance to win up to 50 euros! We will randomly select 19 people among the participants and have a budget of 850 euros for this experiment.

You will not have to log into our website: this personal link will automatically register that you have participated in the experiment.

The experiment will be online only 1 more week, so if you want to have a chance of winning 50 euros, you should make sure to participate in the experiment very soon.

Thank you for your interest in our experiments!

Best regards,

Aurelien Baillon

Appendix B: Experimental Instructions

B.1. Welcome page

Welcome and thank you for taking part in our experiment!

You will be participating in a web-based experiment in economic decision-making. Based on the decisions that you make during the experiment, you might receive a monetary payment up to €50 which will depend on your choices in the experiment.

In the experiment, you will be asked to make seventeen choices distributed among four different tasks. Two of these tasks, each involving one choice, will require you to make decisions that will influence both your own and another participant's outcome. The other two tasks, involving the remaining fifteen choices, concern decisions regarding bets.

At the end of the experiment, when all participants have submitted their answers, for each of the 17 choice situations we will randomly select participants for whom this choice situation will be carried out for real money. For every choice situation we will select different participants. Thus, you have a chance that one choice situation will be carried out for real money for you.

We have attempted to make the instructions of the experiment as clear as possible. However, if you still have trouble understanding a task after reading the instructions, have any other questions, or encounter technical difficulties, please contact XXXXX@ese.eur.nl. You can send an email, close the window, and continue the experiment after you have received a reply.

Click next if you're ready to start the experiment.

Next

B.2. JoD mini-game

In this task, you will be randomly matched with another participant in the experiment. We will refer to this other participant as "Player B". Both you and Player B will receive an endowment of € 25. You have to decide whether to reduce Player B's income or to leave it as it is. If you pay € 1, you can reduce Player B's income by € 10. Player B will be asked to make the same choice regarding your income and will incur the same cost (€ 1) if (s)he chooses to reduce your income.

After Player B and you have decided whether or not to reduce each other's income, a die will be thrown twice. Once for you and another time for Player B.

Let us consider the throw concerning Player B's income. If the die shows 1 or 6 Player B's income will be reduced, independent of your decision. If the die shows any other number (2,3,4,5) then your decision will be realized: If you have decided to reduce Player B's income, the income will be reduced. If you have decided not to reduce Player B's income, the income will not be reduced.

The same procedure will be applied to determine your income: first a throw of a die, then, if the die shows a 1 or a 6, your income will be reduced irrespective of Player B's decision. If the die does not show a 1 or a 6, Player B's decision regarding your income will be carried out.

Please be aware that neither Player B nor you will learn about the outcome of the throws of the die. Therefore, if Player B's income is reduced by € 10, Player B will never learn what the reason for this reduction has been: your decision or the results of the throw of the die. Similarly, if your income is reduced, you will not know whether this is due to Player B's decision or the throw of the die.

Please make your decision: Your endowment in this experiment is € 25.

Do you want to pay € 1 to reduce Player B's income by € 10?

Yes

No

Once you have made your decision, click next.

Next

B.3. Dictator Game

You have been allocated € 50. Your task is to decide how much of this amount to allocate to another individual. The other individual will receive this amount and you will keep the rest.

The other individual will be a randomly selected participant of the experiment. This participant cannot be selected to be paid out for his or her own decisions in the experiment; hence, his or her payoff solely depends on your choice. If you happen to be the randomly selected participant whose choice will be paid out for real, we will make sure that you and the other participants will be invited to receive your payments on different days, so as to rule out any chance that you will meet the other participant. You will not learn the identity of the participant you are matched with, and likewise the other participant cannot learn your identity.

You are now asked to state the amount you wish to allocate to the other participant. This must be a number (integer) between 0 and 50.

Once you have made your decision, click next.

Next

B.4. Ellsberg Tasks

This task involves 9 choices. For each of these choices, one participant will be randomly selected, and his/her decision will be implemented for real, and the resulting outcome will be paid in euros. Please state your decision for each of the following choice tasks.

You will have to pick a colour: red or black, and draw a chip from a bag containing red and black chips. If your colour is drawn you will win €50, but if the other colour is drawn, you will win nothing. You have to decide from which bag you would like to draw a chip: Bag A or Bag B.

- In Bag A, there will be 10 chips. Each chip can only be black or red, but the proportion of each colour will be unknown. The bag will be ready before you choose your colour, but you will not be allowed to check what is in it before choosing a colour and drawing a chip.

- In Bag B, we will put (in front of you) x chips of your colour and 10 - x chips of the other colour.

If x were 0, Bag A would be more interesting because there could be at least one chip with your colour in this bag. If x were 10, Bag B would be more interesting because it would guarantee €50. For x=1, 2, ..., 9, you have to choose the bag from which you want to extract a chip so as to win €50 if you draw a chip of your colour.

	Choice 1 x=1	Choice 2 x=2	Choice 3 x=3	Choice 4 x=4	Choice 5 x=5	Choice 6 x=6	Choice 7 x=7	Choice 8 x=8	Choice 9 x=9
Bag A	unknown proportions of red and black chips	unknown proportions of red and black chips	unknown proportions of red and black chips	unknown proportions of red and black chips	unknown proportions of red and black chips	unknown proportions of red and black chips	unknown proportions of red and black chips	unknown proportions of red and black chips	unknown proportions of red and black chips
Bag B	1 chip of your colour, 9 chips of the other colour	2 chips of your colour, 8 chips of the other colour	3 chips of your colour, 7 chips of the other colour	4 chips of your colour, 6 chips of the other colour	5 chips of your colour, 5 chips of the other colour	6 chips of your colour, 4 chips of the other colour	7 chips of your colour, 3 chips of the other colour	8 chips of your colour, 2 chips of the other colour	9 chips of your colour, 1 chip of the other colour
	A B	A B	A B	A B	A B	A B	A B	A B	A B

Once you have made your decision, click next.

Next

B.5. Compound vs. Simple Lotteries

This task involves 6 choices. For each of these choices, one participant will be randomly selected, and his/her decision will be implemented for real, and the resulting outcome will be paid in euros. Please state your decision for each of the following choice tasks.

Each of the choice tasks involves choosing between an option that involves drawing one chip from a bag and another option that involves drawing multiple chips from a different bag.

In case of drawing multiple chips from the bag, the poker chips you draw will be placed back in the bag and the chips in the bag will be mixed before you extract again, so as to keep the composition of the bag constant. This holds true for all choice situations below.

Please pay attention to both the composition of the bags and the number of extractions, which both vary across tasks.

In each choice situation, you have to choose between two options to win €50.

	Choice 1	Choice 2	Choice 3	Choice 4	Choice 5	Choice 6
Option A	extract 1 time from a bag with 10 red and 10 black chips, win if red	extract 1 time from a bag with 5 red and 15 black chips, win if red	extract 1 time from a bag with 5 red and 15 black chips, win if red	extract 1 time from a bag with 2 red and 18 black chips, win if red	extract 1 time from a bag with 4 red and 16 black chips, win if red	extract 1 time from a bag with 6 red and 14 black chips, win if red
Option B	extract 7 times from a bag with 18 red and 2 black chips, win if 7 times red	extract 5 times from a bag with 15 red and 5 black chips, win if 5 times red	extract 7 times from a bag with 16 red and 4 black chips, win if 7 times red	extract 4 times from a bag with 10 red and 10 black chips, win if 4 times red	extract 6 times from a bag with 15 red and 5 black chips, win if 6 times red	extract 2 times from a bag with 10 red and 10 black chips, win if 2 times red
	A B	A B	A B	A B	A B	A B

Once you have made your decision, click next.

Next

B.6. Confirmation screen and additional questions

Confirmation

Your choices have been registered. Please answer the following questions to validate your participation in the experiment.

Read each of the following statements carefully and indicate how characteristic it is of you according to the following scale:

	Not at all characteristic of me	Slightly characteristic of me	Moderately characteristic of me	Very characteristic of me	Extremely characteristic of me
I worry about what other people will think of me even when I know it doesn't make any difference.					
I am unconcerned even if I know people are forming an unfavorable impression of me.					
I am frequently afraid of other people noticing my shortcomings.					
I rarely worry about what kind of impression I am making on someone.					
I am afraid that others will not approve of me.					
I am afraid that people will find fault with me.					
Other people's opinions of me do not bother me.					
When I am talking to someone, I worry about what they may be thinking about me.					
I am usually worried about what kind of impression I make.					
If I know someone is judging me, it has little effect on me.					
Sometimes I think I am too concerned with what other people think of me.					
I often worry that I will say or do the wrong things.					

Did you use a calculator to make some choices in the experiment?

Yes

No

Please indicate you age, gender, year of study, and nationality.

Age:					
Gender:	Male	Female			
Year of study:	Bachelor 1	Bachelor 2	Bachelor 3	Master	Other
Nationality:	Dutch	Other			
Any comment? (optional)					

B.7. Final screen

Thank you for your participation. Your answers have been recorded.

When the experiment is over, we will let you know whether you have been selected to play one of your choices for real.