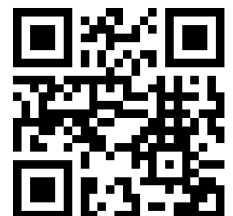working paper

eeecon
[triple:e:con]

©eeecon

The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan

The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

Sebastian Bachler[†], Felix Holzmeister[‡],
Michael Razen[†,*] & Matthias Stefan[†]

[†] Department of Banking and Finance, University of Innsbruck
[‡] Department of Economics, University of Innsbruck

[*] Corresponding author: michael.razen@uibk.ac.at

Abstract

The question of optimal presentation format and choice architecture for investment decisions has gained momentum among researchers, policy makers, and practitioners alike. Motivated by the question how to provide information to investors in a way to improve financial decision-making, we conduct an investment experiment. We implement a 2×2 factorial design to test the effect of presentation format (graphical vs. tabular) and choice architecture/complexity (asset selection vs. predefined portfolios) on decision-making quality. Overall, our results suggest a differential effect of presentation format and choice architecture: Firstly, we find that the graphical presentation format lowers decision-making quality when the environment is more complex (asset selection). Secondly, within graphical presentation, decision quality is higher when the choice architecture is simplified (predefined portfolios). In addition, we find that a simplified choice architecture leads to higher risk-adjusted returns and reduces the relevance of fluid intelligence and numeracy for decision-making quality. *JEL:* G11, G41, G50

Keywords: Investment Decision, Presentation Format, Choice Architecture, Decision-Making Quality, Experimental Finance

1. Introduction

At the core of personal investment decisions is the choice about one's portfolio, which is characterized by a trade-off between risk and return. From behavioral finance, we are well aware that such decisions are difficult (see, e.g., García, 2013), which is why regulators struggle to find the best way to support people in making financial decisions. Among others, these policy decisions include the presentation of relevant information and the design of the decision environment. For instance, the *Markets in Financial Instruments Directive* (*MiFID*) by the European Union requires advisors to provide the relevant information to investors "so that they are reasonably able to understand the nature and risks of the investment service and of the specific type of financial instrument that is being offered and, consequently, to take investment decisions on an informed basis" (Article 19, No. 3). The *Key Investors Information Document*, which regulates investor information in the European Union, specifies this regulation by mandating charts about past fund performance as visual aid.

However, recent experimental evidence shows that it is not *per se* clear which presentation format of relevant information actually improves decision-making. Existing evidence demonstrates that even small changes can have an effect on the perception of relevant information and, in turn, on financial decisions. For instance, Glaser et al. (2019) show that presenting information on past performance in the form of either returns or prices does indeed lead to different estimates of the prospective performance, and Kaufmann et al. (2013) and Bradbury et al. (2015) show that investment decisions vary depending on whether risk is described or whether it is experienced. Similarly, Bateman et al. (2014, 2016) report variation in retirement portfolio decisions with varying presentations of risk, in particular depending on how it is described and whether it is described or displayed graphically (see also Kozup et al., 2008). These findings highlight a tension in using graphs or tabular displays, both of which are abundantly used to present financial information. As noted, many regulations require graphical displays of relevant fund information, although research in other areas indicates that the impact of graphical display of information cannot be generalized as it depends on the task and the type of information (see, e.g., Kelton and Pennington, 2010). In line with the regulation rationale, DeSanctis and Jarvenpaa (1989) provide evidence that financial forecasting is slightly improved by graphical displays.

Yet, existing research also shows that the effectiveness of presentation formats of relevant information cannot be assessed in isolation. Bateman et al. (2014) show that the effect of presentation format of risk on portfolio decisions is moderated by financial literacy. Since the effect of presentation formats on an individual investor's decision also depends on her financial capability and literacy, these dimensions cannot be assessed separately. This insight is highlighted by Foster et al. (2015), showing that presentation formats of information are more effective when investors have the numeracy skills to process the relevant information provided in a way that *should* make it easier to process. In early studies on production decisions, Remus (1984, 1987) argues that both tabular and graphical displays *can* lead to better decisions depending on environmental complexity (see also Kelton and Pennington, 2010). This has also been discussed in the context of *choice*

architectures, where decision environments are systematically modified in order to guide decision-makers towards "improved" decisions (for the individual or for society; Thaler and Sunstein, 2008; Münscher et al., 2016). Thus, not only presentation formats but also the decision environment and individual characteristics are of vital importance for financial decision-making. Since financial literacy varies strongly between people and since it cannot be readily improved in the short run (Fernandes et al., 2014), altering decision situations can provide an important tool to increase financial decision-making quality (Oehler and Wendt, 2017). Along similar lines, Lusardi (2008) argues that investors' decision-making can only be improved by simplifying the decision-making process.

Motivated by the question how to provide information to investors in a way to improve financial decision-making, we combine and contribute to these different strands of literature. The first research question we address in our paper is whether graphical or tabular display of information on investment possibilities – in particular, the corresponding payoff profiles – improve decision-making quality. The second research question is whether lower (higher) investment decision-making complexity increases (decreases) decision-making quality. To answer these two research questions, we conduct an experiment with a 2 (graphical vs. tabular display) × 2 (asset selection vs. choosing from predefined portfolios) factorial treatment design. This way, we cannot only address the two research questions separately, but also examine whether a potential effect of the presentation format interacts with a potential effect of decision-making complexity. This is particularly interesting given the existing results on presentation formats in settings that differ in complexity as described above. Moreover, in light of the prior findings on investors' characteristics potentially mediating the effect of presentation formats, we introduce as a third research question whether participants' cognitive skills (measured as fluid intelligence and numeracy) moderate the impact of the various treatments on decision-making quality.

Overall, our results suggest a differential effect of the treatment variations with respect to the presentation format and the choice architecture on decision-making quality and portfolio performance. First, we find that a graphical presentation format can be detrimental for participants' decision-making quality when they face decision environments involving higher complexity, i.e., when they are asked to build portfolios based on information on the individual assets. Second, within graphical presentation, decision-making quality is higher when the choice architecture is simplified to a selection from predefined portfolio. In addition, we find that a simplified choice architecture – irrespective of whether information is presented graphically or tabularly – leads to a systematic reduction in participants' average portfolio risk while not affecting expected returns. This consequentially translates into a systematic increase in risk-adjusted returns. Finally, we observe that decision-making quality increases with fluid intelligence and numeracy. While the effects of cognitive skills tend to be homogeneous across treatments, we find suggestive evidence that, on the one hand, the adverse effects of lower fluid intelligence are particularly pronounced in treatments in which participants construct portfolios from tabular information. On the other hand, the impact of numeracy tends to be most pronounced when participants are asked to choose from predefined portfolios presented in tabular form. Overall, our findings indicate that simplifying investment decisions reduces the relevance of cognitive skills for decision-making quality, helping those

with relatively lower levels of sophistication to make better decisions. These differential effects of graphical information provision as well as the interaction effects between treatment variations and cognitive abilities observed in our experiment might call for caution when it comes to generalized policy solutions, as tailored solutions for individuals might be the preferable way forward.

2. Experimental Design

Recruitment and data collection. We conducted an online experiment with a total of 325 participants. The experiment was designed using *oTree* (Chen et al., 2016).[1] Participants – bachelor and master students from various fields of study, registered as participants in the *Innsbruck EconLab* – were recruited via HROOT (Bock et al., 2014). After signing-up to take part in the experiment, participants received an individualized link to access the online experiment via e-mail. The experiment was conducted in German and took approximately 20 minutes. Participants received an average payment of €10.18 (sd = €3.01, min = €3.30, max = €18.60).

Allocation decision task. The main task in our experiment is based on the allocation decision task used by Banks et al. (2018) and Holzmeister et al. (2019). It comprises 14 investment decisions, each featuring two different assets, *A* and *B*. The assets were stylized as 50/50-lotteries with state-contingent payoffs. For each of the 14 decisions, participants were asked how they wish to allocate an endowment of €10 to the two assets. Participants' allocation decisions were restricted to investments of integer values summing up to €10 (i.e., participants were required to invest their entire endowment). Thus, for each opportunity set, participants could implement one out of 11 feasible portfolios (0/10, 1/9, 2/8, ..., 10/0). The state of nature (and hence the assets' returns per €1 invested)[2] was determined by a simulated coin toss showing either heads or tails. The state-contingent returns for each asset *A* and *B*, i.e., (h_A, h_B) for heads and (t_A, t_B) for tails, are shown in Table 1. Note that 3 of the 14 decisions (#9, #11, and #14 in Table 1) included a strictly dominating asset (positively sloped opportunity sets).

The order of investment decisions was randomized for each participant to avoid order effects (see, e.g., Carlsson et al., 2012). To prevent cross-decision contamination effects (see, e.g., Cubitt et al., 1998), we implemented a random lottery incentive system, i.e., only one of the 14 investment decisions was chosen at random to determine a participants's final payment (Azrieli et al., 2018).

Decision-making quality measures. The main rationale for the allocation decision task used in our experiment is that it allows quantifying violations of rationality. In particular, following Banks et al. (2018) and Holzmeister et al. (2019), we determine (i) violations of the first-order stochastic dominance principle (*FOSD*; Hadar and Russell, 1969), (ii) violations of state-by-state dominance (*SSD*), and (iii) violations of the generalized axiom of revealed preferences, inspired by the "money

---

[1] Demo versions of the experimental software for each of the four treatments, including all instructions, are available via https://bhrs-de.herokuapp.com/ (in German) and https://bhrs-en.herokuapp.com/ (translated to English).
[2] Note that by "return" we mean the euro payoff of the asset.

Table 1: Payoffs for assets A and B for realizations of heads (h_A, h_B) and tails (t_A, t_B) for each of the 14 opportunity sets. The values correspond to returns (in €) per €1 invested, depending on whether the coin toss shows up heads or tails. For each participant, the order in which the investment decisions were presented has been randomized.

|	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14
h_A	0.38	0.00	1.32	1.37	1.20	0.48	1.27	0.94	1.20	0.57	0.32	1.10	0.92	1.17
t_A	1.44	0.77	0.34	0.24	0.00	0.94	0.09	0.06	0.53	0.59	0.32	0.09	0.76	0.57
h_B	0.58	1.27	0.07	0.50	0.00	1.42	0.41	0.83	0.34	1.50	1.18	0.74	0.94	0.71
t_B	0.18	0.39	0.92	0.82	1.44	0.08	1.16	0.92	0.45	0.22	0.59	0.74	0.62	0.44
Slope	−	−	−	−	−	−	−	−	+	−	+	−	−	+

pump index" (Echenique et al., 2011). In this paper, we propose slightly modified versions of the corresponding measures that are easily applicable to experiments with discrete choices.[3] Moreover, the task emulates the stylized aspects of an investment decision in a concise way, which makes it particularly suitable for the use in an experimental setting.

Experimental treatments. Upon accessing the online experiment, participants were randomly assigned to one treatment of a 2×2 factorial between-subjects design. Along the first treatment dimension, we varied the way how the information on the state-contingent payoffs (of either individual assets or portfolios) was presented to participants: in "Tabular" or "Graphical" form (the latter by means of bar charts without data labels). Along the second dimension, we varied whether participants were informed about individual assets' state-contingent payoffs or whether participants were informed about the state-contingent payoffs of the set of feasible portfolios based on the individual assets.[4] Thus, based on the necessary steps to come up with a portfolio choice, the former situation represents a decision of higher complexity than the latter. We label these treatment arms as "Asset" and "Portfolio," respectively. Figure 1 illustrates the treatment design. Figures A1–A4 in the Appendix show screenshots of the experimental implementation of the decision environment for one opportunity set across the four treatment conditions.

For ease of exposition, we identify the treatments by T ("Tabular") and G ("Graphical") to denote the representation of information, and by A ("Asset") and P ("Portfolio") to denote the choice architecture. Thus, T indicates treatments in which the information is presented in tabular format, whereas G indicates the treatments in which information is presented graphically. Likewise, A

---

[3] Note that, by the design of our experiment, violations of the rationality principles are less likely if the decision-maker is risk neutral, as this would render diversification irrelevant. In our stylized environment, the two assets span a straight line as the set of feasible portfolios in the state-contingent payoff space (see Figure B1 in the Appendix). More precisely, due to the discrete portfolio weights, the set of feasible portfolios is a set of equidistant points on that line. Consider a risk neutral decision-maker: Choosing a portfolio is a straightforward exercise as she will invest the entire endowment in the asset yielding the highest expected return. On the other hand, consider a highly risk averse decision-maker: In order to construct well-balanced portfolios which account for her level of risk aversion, the decision-maker will – for downward-sloping opportunity sets – choose portfolios in the interior of the opportunity sets (rather than boundary portfolios), which increases the likelihood of violating the principle of first-order stochastic dominance (*FOSD*), state-by-state dominance (*SSD*), and the generalized axiom of revealed preferences (*GARP*).

[4] The set of predefined portfolios covered the feasible allocations (i.e., 0/10, 1/9, 2/8, . . . , 10/0).
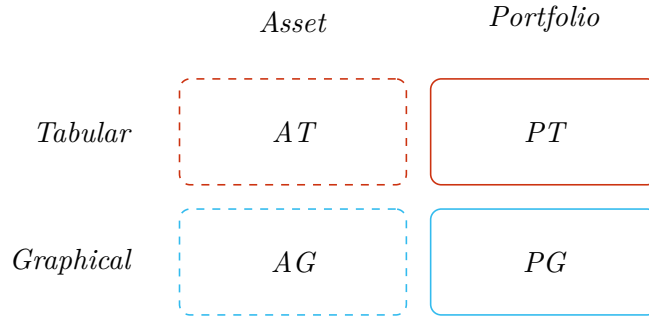
Figure 1: Treatment Design. The figure illustrates the 2×2 factorial treatment design implemented in the experiment, alongside the treatment labeling and color coding as used throughout the text.

indicates the treatments in which participants obtain information about the individual assets, whereas P indicates the treatments in which participants obtain information about predefined portfolios. In concatenated form, AT, PT, AG, and PG each refer to one specific cell in the treatment design. PT, for instance, denotes the treatment in which participants have to choose from predefined portfolios presented in tabular format.

Elicitation of cognitive skills, decision self-assessments, and risk preferences. To obtain validated measures for fluid intelligence and numeracy skills, we presented participants with a subset of eight items of the *Advanced Raven Matrices* (Raven, 2000) and the numeracy test by Weller et al. (2013).[5] These two additional tasks were incentivized with a piece-rate of €0.30 per correct item, i.e., participants could earn up to €2.40 in each of the two tasks. After the allocation decisions task, participants were asked four self-assessment questions. The first three of them were on the perceived riskiness of the decisions made, the difficulty of making decisions that match one's preferences, and the satisfaction with the decisions made. The fourth item asked participants to self-assess their general willingness to take risk as in Dohmen et al. (2011). Each of these four questions had to be answered on a 7-point Likert scale.[6]

3. Results

Summary statistics. Table 2 provides descriptive statistics on the outcomes of the portfolio choices and on cognitive skills, decision self-assessments, and risk preferences. The results on decision-making quality reveal interesting differences in the severity of violations of the three decision-making quality principles. Missing out on the advantages of state-by-state dominance

---

[5] To avoid potential task recognition, we replaced the cognitive reflection questions from Frederick (2005) in this test by the ones suggested by Toplak et al. (2014).

[6] For details on the elicitation of cognitive skills and self-assessments, please refer to the demo versions of our experimental software in German (https://bhrs-de.herokuapp.com/) or English (https://bhrs-en.herokuapp.com/).

(*SSD*) costs participants on average €0.77 per investment decision. In contrast, violations of less apparent first-order stochastic dominance (*FOSD*) and violations of the generalized axiom of revealed preferences (*GARP*) lead to average losses of €0.12 and €0.14 per decision, respectively (note that our measure *GARP* is based on the comparison of two decisions, so the resulting expected loss needs to be divided by two to obtain a per-decision value). For the summary statistics, we put these numbers into further context by relating them to (i) the corresponding expected payoffs of the chosen portfolio and (ii) the maximum potential loss for each decision and quality criterion. First, (i) tells us the %-loss in expected returns due to suboptimal decisions. We find that, on average, *FOSD* amounts to 2.01%, *SSD* to 12.60%, and *GARP* to 4.16% of the expected payoff of the chosen portfolio. Second, (ii) allows for a better cross-principle comparison, as it controls for task-specific differences in the potential scope for error. Our results show that, on average, *FOSD* amounts to 4.21%, *SSD* to 17.13%, and *GARP* to 17.43% of the maximum potential loss of the respective decision and decision-making quality measure, revealing also considerable variation in the decision-making quality measures.

Table 2: Summary statistics ($n = 325$) for the three measures of decision-making quality (*FOSD*: violations of first-order stochastic dominance; *SSD*: violations of state-by-state dominance; *GARP*: violations of the generalized axiom of revealed preferences), the three payoff profile or performance indicators (*ER*: mean expected portfolio return; *SD*: mean portfolio standard deviation; *CV*: mean portfolio variation coefficient), participants' cognitive skills and risk preferences (*APM*: fluid intelligence; *NUM*: numeracy; *RT*: self-reported risk tolerance), and participants' self-assessments (perceived risk; perceived difficulty; satisfaction). Statistics refer to the average of the 14 decisions per participant.

	Mean	*St. Dev.*	*Min*	*Median*	*Max*
FOSD	0.12	0.17	0.00	0.06	1.22
SSD	0.77	0.84	0.00	0.44	3.83
GARP	0.27	0.26	0.00	0.20	1.78
ER	7.49	0.41	6.33	7.50	8.26
SD	2.24	0.73	0.63	2.09	4.77
CV	0.30	0.09	0.09	0.28	0.64
APM	4.97	1.71	0.00	5.00	8.00
NUM	4.99	1.65	0.00	5.00	8.00
RT	3.99	1.41	1.00	4.00	7.00
Perc. Risk	3.81	1.60	1.00	4.00	7.00
Perc. Diff	3.56	1.56	1.00	3.00	7.00
Satisfaction	4.81	1.29	1.00	5.00	7.00

Similarly, the maximum losses participants incurred (averaged across all of their 14 decisions) were €1.22 due to *FOSD*, €3.83 due to *SSD*, and €1.78 due to *GARP*, or 19.19%, 56.31%, and 28.12% of the corresponding expected payoff, respectively. This indicates that there is indeed substantial room for improvement in investment decisions. Note that it was feasible to avoid errors in decision-making in our setup, as shown by the fact that, for all three measures, the minimum loss is €0.00. As the corresponding values show averages, a score of 0 requires to avoid errors in all 14 decisions.

For the statistical analyses, we z-standardize all dependent variables (i.e., the three decision-making quality measures and the three performance indicators). Participants' numeracy (NUM) and fluid intelligence scores (APM) enter all analyses in levels, i.e., on a scale from 0 to 8. The estimated coefficients thus represent effect sizes in terms of standard deviations for a one unit change in covariates. Participants' self-reported risk attitudes are treated as ordinal variables to allow for potentially non-linear effects; accordingly, we treat all possible values on the scale as dichotomous variables.

Observation 1. *Presenting information graphically (G) impairs decision-making quality if participants obtain information on the individual assets (A), but has no systematic effect if participants choose from predefined portfolios (P). Moreover, choosing from predefined portfolios (P) significantly improves decision-making quality if information is presented graphically (G), but has no systematic effect if asset returns are presented tabularly (T).*

Support: Figure 2 shows interaction plots based on ordinary least squares regressions of (a) violations of first-order stochastic dominance ($FOSD$), (b) violations of state-by-state dominance (SSD), and (c) violations of the generalized axiom of revealed preferences ($GARP$) on treatment indicators for "Portfolio" (P) and "Graphical" (G), and their interaction term ($P{\times}G$). The corresponding regression estimates are provided in Table C1 in the Appendix (note that the base categories are "Asset" (A) and "Tabular" (T)).
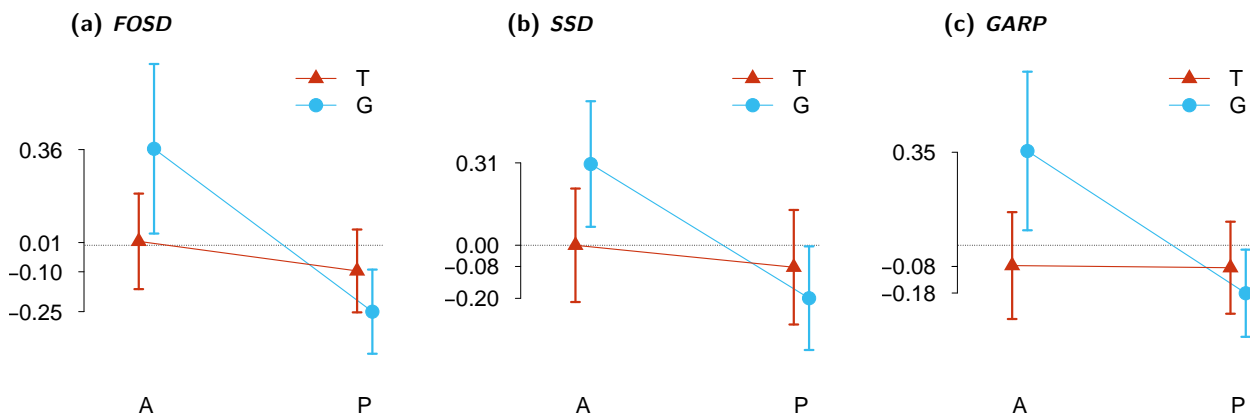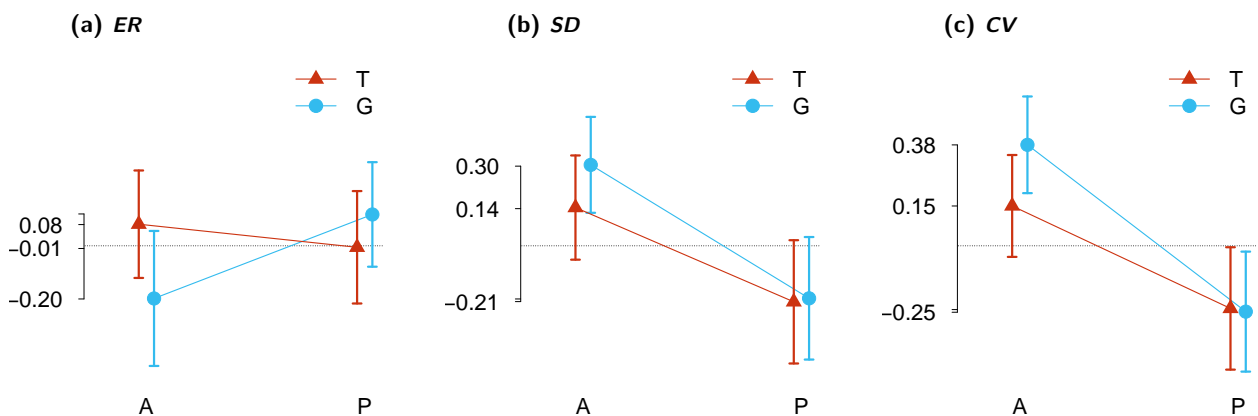

Figure 2: Interaction plots of ordinary least squares regressions of **(a)** violations of first-order stochastic dominance ($FOSD$), **(b)** violations of state-by-state dominance (SSD), and **(c)** violations of the generalized axiom of revealed preferences ($GARP$) on indicators for the treatments "Portfolio" (P) and "Graphical" (G), and their interaction term ($P{\times}G$). The three dependent variables are z-standardized scores. The dotted horizontal line indicates zero. Error bars indicate 95% confidence intervals based on robust standard errors. The regression estimates are provided in Table C1 in the Appendix.

For all three measures of rationality violations, treatment effects are homogeneous. While the effect of offering predefined portfolios (P) is insignificant if the information is presented in tabular format (T), it turns out to be significantly negative (indicating an improvement in decision-making quality) when information is presented graphically ($P + P{\times}G$; $FOSD$: $b = -0.613$, $p < 0.001$; SSD:

$b = -0.505$, $p = 0.001$; and *GARP*: $b = -0.536$, $p = 0.002$). Providing graphical information on the individual assets (G) tends to increase the magnitude of rationality violations if participants are required to allocate their endowment to the two assets (*FOSD*: $b = 0.349$, $p = 0.062$; *SSD*: $b = 0.306$, $p = 0.060$; and *GARP*: $b = 0.431$, $p = 0.019$). Due to the sizeable negative interaction effect of "Portfolio" and "Graphical" ($P{\times}G$), however, the effect of presenting information graphically rather than tabularly for predefined portfolios ($G + P{\times}G$) turns out being insignificant for all three measures of decision-making quality. Notably, the results on *FOSD* and *SSD* are robust to analyses on the non-aggregated level, i.e., on a per-decision basis, estimated using random effects regressions (see Table C2 for details). The results reported for *GARP* cannot be estimated on a per-decision level since the measure is based on violations of the generalized axiom of revealed preferences and thus by definition requires pairwise comparisons of decisions (see Figure B2 in the Appendix).

Observation 2. *Presentation format and choice architecture do not have a systematic effect on mean expected portfolio returns. However, participants take significantly less risk when they choose from predefined portfolios (P), irrespective of whether information is presented in tabular (T) or graphical format (G). These two findings translate into significantly higher risk-adjusted returns, i.e., smaller variation coefficients, if participants can choose from a set of predefined portfolios, regardless of whether information is presented tabularly or graphically.*

Support: Figure 3 shows interaction plots based on ordinary least squares regressions of (a) the mean expected portfolio return (*ER*), (b) the mean portfolio standard deviation (*SD*), and (c) the mean portfolio variation coefficient (*CV*) on treatment indicators for "Portfolio" (P) and "Graphical" (G), and their interaction term ($P{\times}G$). The corresponding regression estimates are provided in Table C3 in the Appendix.

Figure 3: Interaction plots of ordinary least squares regressions of **(a)** the mean expected portfolio return (*ER*), **(b)** the mean portfolio standard deviation (*SD*), and **(c)** the mean portfolio variation coefficient (*CV*) on indicators for the treatments "Portfolio" (P) and "Graphical" (G), and their interaction term ($P{\times}G$). The three dependent variables are z-standardized scores. The dotted horizontal line indicates zero. Error bars indicate 95% confidence intervals based on robust standard errors. The regression estimates are provided in Table C3 in the Appendix.

With respect to mean expected portfolio returns (ER), neither choice architecture (P) and presentation format (G), nor the interaction effect ($P{\times}G$) are statistically significant, i.e., we do not find evidence that the treatment interventions have a systematic effect on expected returns. However, we find that the mean standard deviation of the portfolios (SD) is significantly lower in treatments in which participants choose from predefined portfolios (P: $b = -0.355$, $p = 0.022$; $P + P{\times}G = b = -0.502$, $p < 0.001$). These two findings translate into significantly higher risk-adjusted returns, i.e., smaller coefficients of variation (P: $b = -0.386$, $p = 0.012$; $P + P{\times}G = b = -0.628$, $p < 0.001$). As for the measures of decision-making quality, the results on ER, SD, and CV turn out to be robust in analyses on the per-decision level, estimated using random effects regressions (see Table C4 for details).

Observation 3. *Participants who score higher on fluid intelligence (APM) and numeracy (NUM) exhibit, on average, higher decision-making quality. While we do not observe significant interactions between cognitive skills and treatments, our results suggest that the positive effects are particularly pronounced for fluid intelligence when participants have to construct portfolios from tabular information on the assets (PT), and for numeracy when participants can choose from predefined portfolios presented in tabular form (AT).*

Support: Figure 4 shows the average effects of a unit increase in *APM* (upper panel) and *NUM* (lower panel) on the three (z-standardized) measures of decision-making quality. As *FOSD*, *SSD*, and *GARP* represent avoidable losses, negative signs indicate improvements in decision-making quality. Our results suggest that both fluid intelligence and numeracy support participants in the decision-making process (all panels). In economic terms, scoring one unit higher in *APM*, pooled across all treatments, reduces *FOSD*-errors by -0.130 standard deviations ($p < 0.001$), *SSD*-errors by -0.122 standard deviations ($p < 0.001$), and *GARP* by -0.115 standard deviations ($p < 0.001$). Likewise, a one-unit increase in *NUM*, pooled across all treatments, is associated with a decrease in *FOSD* by -0.101 standard deviations ($p = 0.013$), a decrease in *SSD* by -0.129 standard deviations ($p < 0.001$), and a decrease of *GARP* by -0.088 standard deviations ($p = 0.022$).[7]

As illustrated by Figure 4, the (consistently advantageous) effects of *APM* and *NUM* on decision-making quality do vary in size across the four treatments, but most of the times the estimates fall in similar ranges. Due to the relatively small sample size of approximately 80 participants per treatment, of course, confidence intervals tend to be lucidly inflated. In terms of statistical significance, the effects are most pronounced for fluid intelligence in treatment *PT*, and for numeracy in treatment *AT*. Pairwise comparisons, however, reveal no significant differences in the effects of fluid intelligence and numeracy on decision-making quality across treatments (see the corresponding regression estimates in Table C5 for details).

---

[7] The coefficient estimates reported for *APM* and *NUM* are based on ordinary least squares regressions (with robust standard errors) of participant-level means of (i) *FOSD*, (ii) *SSD*, (iii) *GARP* on participants' fluid intelligence score (*APM*) and numeracy score (*NUM*), controlling for individual-level risk preferences. $n = 325$ in all three regressions; adjusted R^2s vary between 0.099 and 0.124, i.e., approximately 10% of the variation in the three decision-making quality measures (on the participant level) is explained by the variation in participants' fluid intelligence and numeracy.
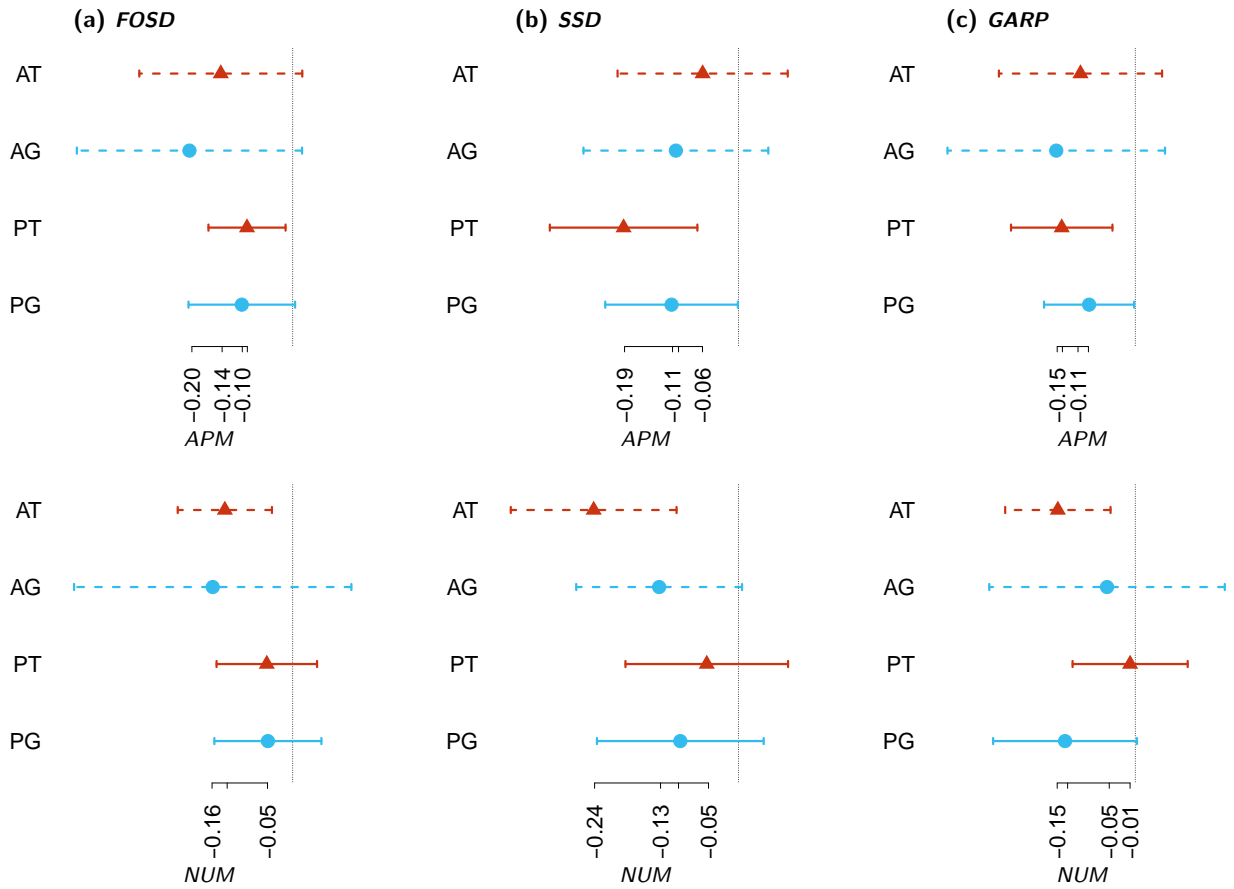
Figure 4: Coefficient plots of fluid intelligence (*APM*; top row) and numeracy (*NUM*; bottom row) conditional on the different treatments as based on ordinary least squares regressions of **(a)** violations of first-order stochastic dominance (*FOSD*), **(b)** violations of state-by-state dominance (*SSD*), and **(c)** violations of the generalized axiom of revealed preferences (*GARP*) on treatment indicators for "Portfolio" (*P*) and "Graphical" (*G*), participants fluid intelligence (*APM*) and numeracy (*NUM*) scores, and the relevant two- and three-way interaction terms (controlling for participants' self-reported risk tolerance). The three dependent variables are z-standardized scores. The dotted vertical line indicates zero. Error bars indicate 95% confidence intervals based on robust standard errors. The regression estimates are provided in Table C5 in the Appendix.

To obtain a more differentiated picture, Figure 5 also includes level effects across treatments and depicts the resulting linear predictions for the entire range of *APM* and *NUM* scores (the predicted slopes are the respective coefficient estimates shown in Figure 4). The prediction plots also illustrate some of our previous findings. Among all treatments, having to construct portfolios based on graphical information (*AG*) produces the largest avoidable losses. Moreover, the consistently downward sloping predictions highlight that cognitive skills reduce errors in decision-making.

Note that policy-relevant interaction effects between cognitive skills and treatments will translate, graphically, to intersecting predictions in Figure 5: If the predictions intersect, this indicates that optimal presentation format and choice architecture might differ depending on an investor's cognitive skills. In such cases, the interaction effect between treatment and cognitive skills outweighs the

corresponding level effects. In Figure 5, we see that the above-mentioned interaction effects indeed lead to intersecting predictions. While the detrimental effect of treatment AT (see also Observation 1) prevails across all levels of fluid intelligence and numeracy, the "second-worst" combination of presentation format and choice architecture seems to indeed depend on the investors' cognitive skills. The upper panel and lower panel confirm the more pronounced effects of fluid intelligence in treatment PT and of numeracy in treatment AT, respectively. The latter observation is even stronger and suggests that having to construct portfolios is particularly disadvantageous for investors with lower numeracy scores, regardless of whether the information is presented graphically or tabularly.

Overall, our findings indicate that simplifying investment decisions by letting participants choose from portfolios instead of individual assets reduces the relevance of APM and NUM for decision-making quality. Consequently, providing investors with easier choice architectures can help those with relatively lower fluid intelligence or numeracy to make better decisions.

Observation 4. *Participants who score higher on fluid intelligence (APM) and numeracy (NUM) choose, on average, asset allocations with higher expected returns. The effect of numeracy is particularly strong when participants have to construct portfolios based on tabular information (AT). While higher numeracy tends to come with higher risk taking, we observe no consistent effect for fluid intelligence. Aside from the above effect, we find no significant interactions between cognitive skills and presentation formats.*

Support: Figure 6 shows the average effects of a one-standard-deviation increase in APM (upper panel) and NUM (lower panel) on the three (z-standardized) performance indicators (see Table C6 for the corresponding regression analyses). Pooled across all four treatments, both APM and NUM have a significantly positive effect on mean expected portfolio returns. On average, a one-unit increase in APM induces an increase in the mean expected return of 0.120 standard deviations ($p < 0.001$), whereas a one-unit increase in NUM is associated with an increase of 0.136 standard deviations ($p < 0.001$).[8] Column (a) in Figure 6 indicates that the effects of both fluid intelligence and numeracy tend to be consistent across all treatments: Pairwise comparisons reveal no significant interactions between cognitive skills and treatments, except for the difference in the effect of numeracy between AT and PT.

On the one hand, we find no systematic effect of fluid intelligence on risk taking, neither when pooling all four treatments ($b = -0.013$, $p = 0.711$) nor when considering the impact for each treatment separately (column (b) in Figure 6, upper panel). On the other hand, numeracy does positively correlate with risk taking. This result is consistent with other findings in the literature. For instance, Lusardi and Mitchell (2011) find that numeracy is related to higher stock market participation, particularly for those with higher education. Similarly, Almenberg and Widmark

---

[8] Estimates referring to pooled models are based on ordinary least squares regression (with robust standard errors) of (i) the mean expected portfolio return (ER), (ii) the mean portfolio standard deviation (SD), and (iii) the mean variation coefficient (CV) on fluid intelligence (APM) and numeracy (NUM), pooled across all four treatments. All three regressions control for participants' self-reported risk attitudes. $n = 325$ in all three regressions; adjusted R^2s are 0.156 for ER, 0.098 for SD, and 0.087 for CV.
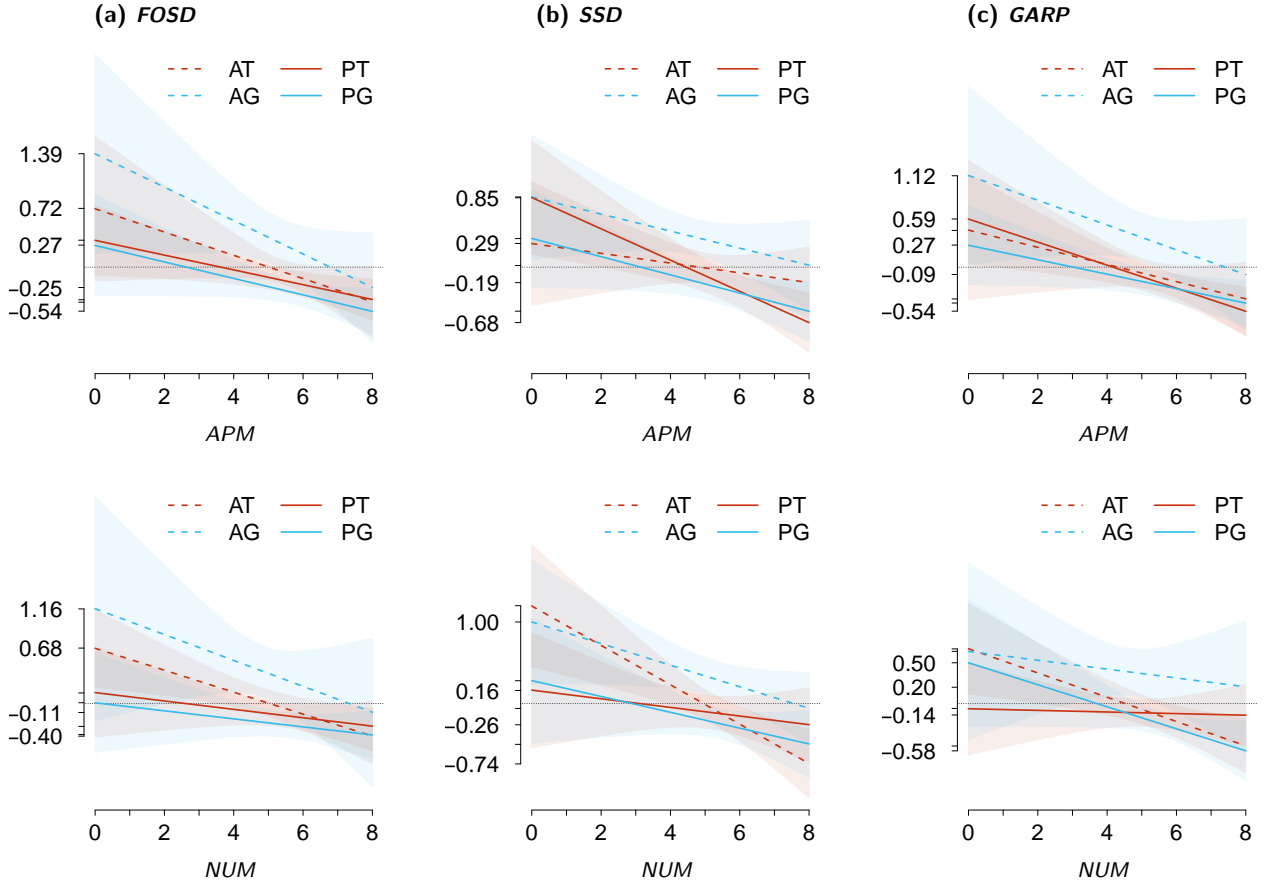
Figure 5: Linear predictions of **(a)** violations of first-order stochastic dominance (*FOSD*), **(b)** violations of state-by-state dominance (*SSD*), and **(c)** violations of the generalized axiom of revealed preferences (*GARP*), conditional on the variation in fluid intelligence (*APM*; top row) and numeracy (*NUM*; bottom row), separated for the four treatment conditions based on ordinary least squares regressions for the treatment indicators for "Portfolio" (*P*) and "Graphical" (*G*), participants fluid intelligence (*APM*) and numeracy (*NUM*) scores, and the relevant two- and three-way interaction terms (controlling for participants' self-reported risk tolerance). The three dependent variables are *z*-standardized scores. The black horizontal line indicates zero. Shaded areas indicate 95% confidence intervals based on robust standard errors. The regression estimates are provided in Table C5 in the Appendix.

(2011) find that numeracy is positively correlated with participation in both stock and housing markets. Pooled across all treatments, the effect of numeracy on portfolio risk (*SD*) turns out to be significantly positive: For a one-unit increase in *NUM*, on average, participants' mean portfolio standard deviation increases by 0.088 standard deviations ($p = 0.004$). This effect tends to be relatively consistent across three of four treatment conditions (column (b) in Figure 6, lower panel). Only in treatment *AG*, portfolio risk and numeracy seem to be unrelated.

Interestingly, pooled across all treatments, the effect of fluid intelligence on risk-adjusted returns turns out to be small and insignificant ($b = -0.042$, $p = 0.234$). As illustrated by the upper panel in column (c) in Figure 6, the effect of *APM* on *CV* tends to vary across the different conditions, but does not reveal systematic differences attributable to the treatment variations.
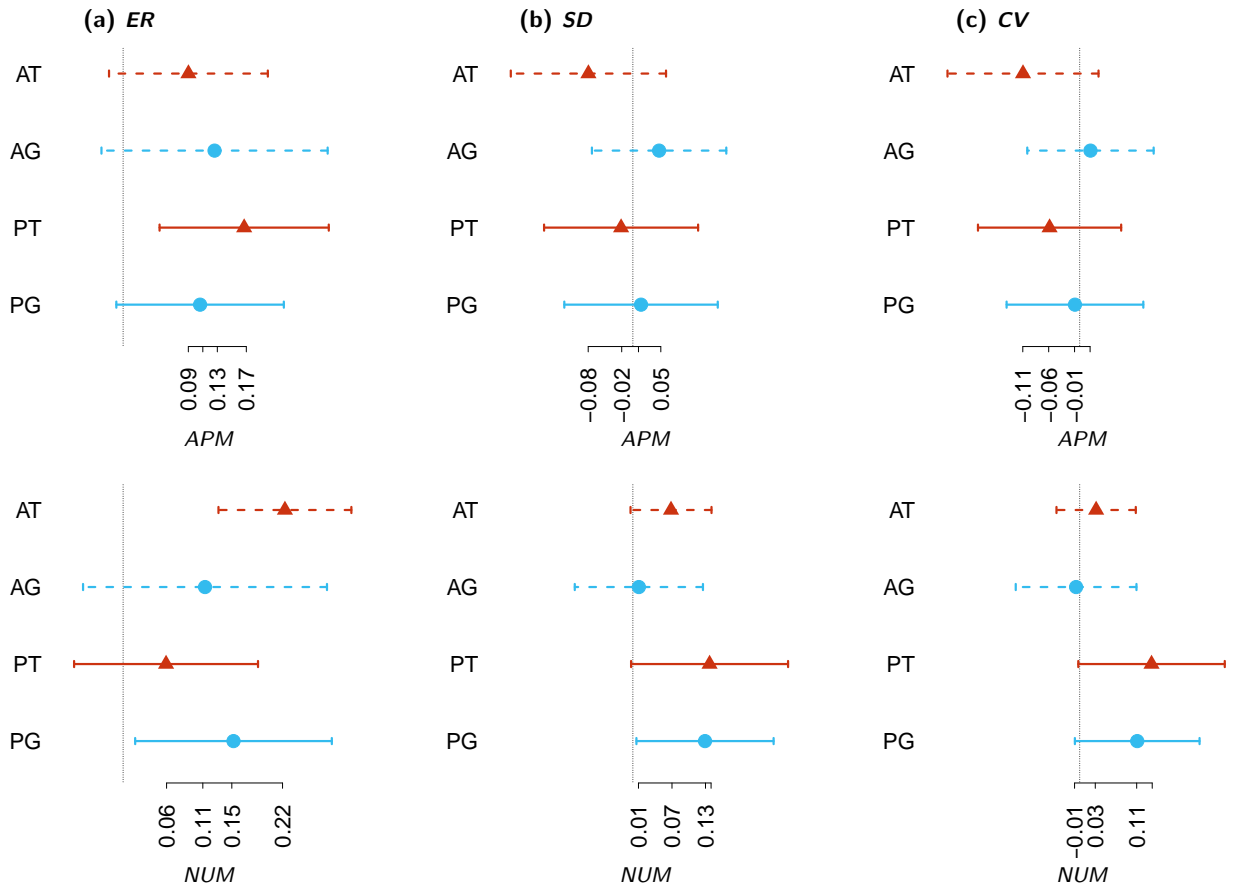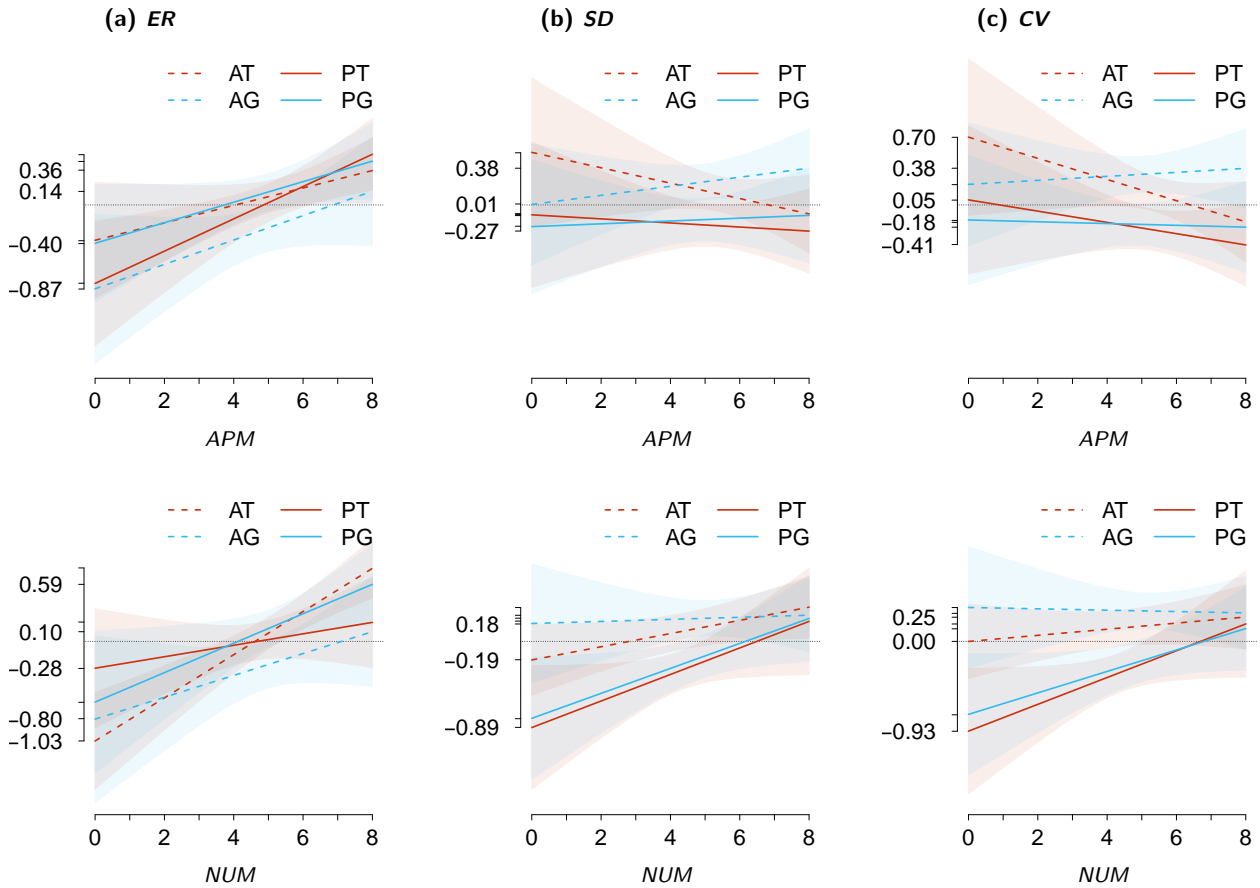
Figure 6: Coefficient plots of fluid intelligence (*APM*; top row) and numeracy (*NUM*; bottom row) conditional on the different treatments as based on ordinary least squares regressions of **(a)** the mean expected portfolio return (*ER*), **(b)** the mean portfolio standard deviation (*SD*), and **(c)** the mean portfolio variation coefficient (*CV*) on treatment indicators for "Portfolio" (*P*) and "Graphical" (*G*), participants fluid intelligence (*APM*) and numeracy (*NUM*) scores, and the relevant two- and three-way interaction terms (controlling for participants' self-reported risk tolerance). The three dependent variables are *z*-standardized scores. The dotted vertical line indicates zero. Error bars indicate 95% confidence intervals based on robust standard errors. The regression estimates are provided in Table C6 in the Appendix.

Numeracy, however, tends to decrease risk-adjusted returns when pooling over all four treatments: On average, a one-unit increase in numeracy increases the mean variation coefficient by 0.070 ($p = 0.023$). As illustrated by the lower panel in column (c) in Figure 6, the negative effect of numeracy on risk-adjusted returns is driven by those treatments in which participants choose from predefined portfolios (treatments *PT* and *PG*).

As an alternative means of representation of the results, we illustrate the effects of fluid intelligence (*APM*) and numeracy (*NUM*) on the performance indicators in terms of linear predictions in Figure 7. Column (a) confirms the positive effects of both fluid intelligence and numeracy on expected returns. As for numeracy, the lower graph also highlights the distinct role of treatment *AT* (building a portfolio based on tabular information). Among all treatments, it leads to the lowest expected returns for participants with low numeracy scores, but to the highest expected returns for

participants with high numeracy scores. As for risk taking, column (b) shows the positive correlation between numeracy and risk taking, but, at the same time, that the effect is moderated when participants have to build the portfolio themselves (lower panel). Finally, column (c) illustrates that the above-mentioned adverse effect of numeracy on risk-adjusted returns (i.e., the positive effect on CV) is largely driven by the treatments where participants choose from pre-defined portfolios (P).

Figure 7: Linear predictions of **(a)** the mean expected portfolio return (ER), **(b)** the mean portfolio standard deviation (SD), and **(c)** the mean portfolio variation coefficient (CV), conditional on the variation in fluid intelligence (APM; top row) and numeracy (NUM; bottom row), separated for the four treatment conditions based on ordinary least squares regressions for the treatment indicators for "Portfolio" (P) and "Graphical" (G), participants fluid intelligence (APM) and numeracy (NUM) scores, and the relevant two- and three-way interaction terms (controlling for participants' self-reported risk tolerance). The three dependent variables are z-standardized scores. The black horizontal line indicates zero. Shaded areas indicate 95% confidence intervals based on robust standard errors. The regression estimates are provided in Table C6 in the Appendix.

4. Discussion and Conclusion

In this paper, we reported from an online experiment based on a 2×2 factorial treatment design, varying the presentation format (graphical vs. tabular) and the complexity of the choice architecture (asset selection vs. choice from predefined portfolios). Our results showed that these two parameters do indeed play a role for decision-making quality of investment allocation choices. Of all treatments, having to construct one's own portfolio based on graphical information about the assets at hand led to the lowest decision-making quality scores. However, this does not necessarily mean that graphical representation was disadvantageous, as the challenge of asset weighting seems to be a crucial factor. When making a decision based on graphical presentation, a simplification of the choice architecture, i.e., allowing participants to choose from a set of predefined portfolios, increases decision-making quality. Moreover, a simplified choice architecture also helped participants to increase their risk-adjusted returns. Finally, in our experiment, decision-making quality increased with fluid intelligence and numeracy. We also found suggestive evidence that constructing portfolios from tabular information was particularly hard for participants with relatively lower fluid intelligence scores, while choosing from predefined portfolios presented in a tabular form was disadvantageous for investors with relatively lower numeracy scores. Interestingly, we found this tendency in a rather homogeneous and already well-educated sample of university students. Overall, these findings indicate that a simplified investment decision environment reduces the impact of cognitive skills on decision-making quality, helping those with relatively lower levels of sophistication to make better decisions.

We consider our results to be policy-relevant for at least two reasons: First, they show that the choice architecture of a decision environment can indeed have beneficial or detrimental effects on decision-making quality. This indicates, as Lusardi (2008) points out with regards to household savings, that in order to improve individuals' financial choices the decision-making process needs to be simplified and barriers for processing information should be reduced (see also Münscher et al., 2016). Lusardi (2008) provides the example of narrowing the number of options available in order to increase saving. In line, we find that a simplified choice architecture can increase the quality of observed choices in our laboratory portfolio decisions and support those with lower fluid intelligence and numeracy. Second, our results indicate that there are also interaction effects of presentation format, choice architecture, and cognitive skills. This highlights that one-size-fits-all solutions will often come at the sacrifice of decision-making quality, as the underlying principles might not be generalizable to all decision situations and individuals. In contrast, accounting for investor heterogeneity when designing information presentation formats and choice architecture can further support decision-making quality.

References

Almenberg, J., & Widmark, O. (2011). Numeracy, financial literacy and participation in asset markets. *European Economics: Labor & Social Conditions eJournal.*

Azrieli, Y., Chambers, C. P., & Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, *126*(4), 1472–1503.

Banerjee, S., & Murphy, J. (2009). A simplified test for preference rationality of two-commodity choice. *Experimental Economics*, *12*(2), 252–252.

Banks, J., Carvalho, L., & Perez-Arce, F. (2018). Education, decision-making, and economic rationality. *Review of Economics and Statistics*, *101*(3), 428–441.

Bateman, H., Eckert, C., Geweke, J., Louviere, J., Satchell, S., & Thorp, S. (2014). Financial competence, risk presentation and retirement portfolio preferences. *Journal of Pension Economics and Finance*, *13*(1), 27–61.

Bateman, H., Eckert, C., Geweke, J., Louviere, J., Satchell, S., & Thorp, S. (2016). Risk Presentation and Portfolio Choice. *Review of Finance*, *20*(1), 201–229.

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, *71*, 117–120.

Bradbury, M., Hens, T., & Zeisberger, S. (2015). Improving Investment Decisions with Simulated Experience. *Review of Finance*, *19*, 1019–1052.

Carlsson, F., Mørkbak, M. R., & Olsen, S. B. (2012). The first time is the hardest: A test of ordering effects in choice experiments. *Journal of Choice Modelling*, *5*(2), 19–37.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, *1*, 115–131.

DeSanctis, G., & Jarvenpaa, S. (1989). Graphical presentation of accounting data for financial forecasting: An experimental investigation. *Accounting, Organizations and Society*, *14*(5–6), 509–525.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550.

Echenique, F., Lee, S., & Shum, M. (2011). The Money Pump as a measure of revealed preference violations. *Journal of Political Economy*, *119*(6), 1201–1223.

Fernandes, D., Lynch, J., & Netemeyer, R. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*, *60*(8), 1861–2109.

Foster, D., Ng, J., & Wee, M. (2015). Presentation Format and Financial Literacy: Accessibility and Assessability of Retirement Savings Statements. *The Journal of Consumer Affairs*, *49*(3), 519–549.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

García, M. (2013). Financial Education and Behavioral Finance: New Insights into the Role of Information in Financial Decisions. *Journal of Economic Surveys*, *27*(2), 297–315.

Glaser, M., Iliewa, Z., & Weber, M. (2019). Thinking about Prices versus Thinking about Returns in Financial Markets. *The Journal of Finance*, *74*(6), 2997–3039.

Hadar, J., & Russell, W. (1969). Rules for ordering uncertain prospects. *American Economic Review*, *59*(1), 25–34.

Holzmeister, F., Holmén, M., Kirchler, M., Stefan, M., & Wengström, E. (2019). Delegated decision-making in Finance. *Working Papers in Economics and Statistics 2019-21, University of Innsbruck.*

Kaufmann, C., Weber, M., & Haisley, E. (2013). The Role of Experience Sampling and Graphical Displays on One's Investment Risk Appetite. *Management Science*, *59*(2), 323–340.

Kelton, A., & Pennington, R. (2010). The Effects of Information Presentation Format on Judgment and Decision Making: A Review of the Information Systems Research. *Journal of Information Systems*, *24*(2), 79–105.

Kozup, J., Howlett, E., & Pagano, M. (2008). The Effects of Summary Information on Consumer Perceptions of Mutual Fund Characteristics. *The Journal of Consumer Affairs*, *42*(1), 37–59.

Lusardi, A. (2008). Household saving behavior: the role of financial literacy, information, and financial education programs. *NBER Working Paper*, *1382*.

Lusardi, A., & Mitchell, O. S. (2011). *Financial literacy and planning: Implications for retirement wellbeing.* (M. O. S. & L. A., Eds.). Oxford, Oxford University Press.

Münscher, R., Vetter, M., & Scheuerle, T. (2016). A Review and Taxonomy of Choice Architecture Techniques. *Journal of Behavioral Decision Making*, *29*, 511–524.

Oehler, S., & Wendt, S. (2017). Good Consumer Information: the Information Paradigm at its (Dead) End? *Journal of Consumer Policy*, *40*, 179–191.

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1–48.

Remus, W. (1984). An Empirical Investigation of the Impact of Graphical and Tabular Data Presentations on Decision Making. *Management Science*, *30*(5), 533–542.

Remus, W. (1987). A Study of Graphical and Tabular Displays and Their Interaction with Environmental Complexity. *Management Science*, *33*(9), 1200–1204.

Thaler, R., & Sunstein, C. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness.* New Haven, Yale University Press.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking and Reasoning, 20*(2), 147–168.

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making, 26*(2), 198–212.

Appendices

for Online Publication

The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

Sebastian Bachler[†], Felix Holzmeister[‡],
Michael Razen[†,*] & Matthias Stefan[†]

[†] Department of Banking and Finance, University of Innsbruck
[‡] Department of Economics, University of Innsbruck

[*] Corresponding author: michael.razen@uibk.ac.at

Contents

A. Screenshots of the Decision Environment

Figures A1–A4 show screenshots of the experimental implementation of one particular choice set in the four treatment conditions. In all four illustrations, asset A pays €1.00 if the coin shows heads, and €0.40 if the coin shows tails; asset B yields a payoff of €0.5 if the coin shows heads, and €0.80 if the coin shows tails.
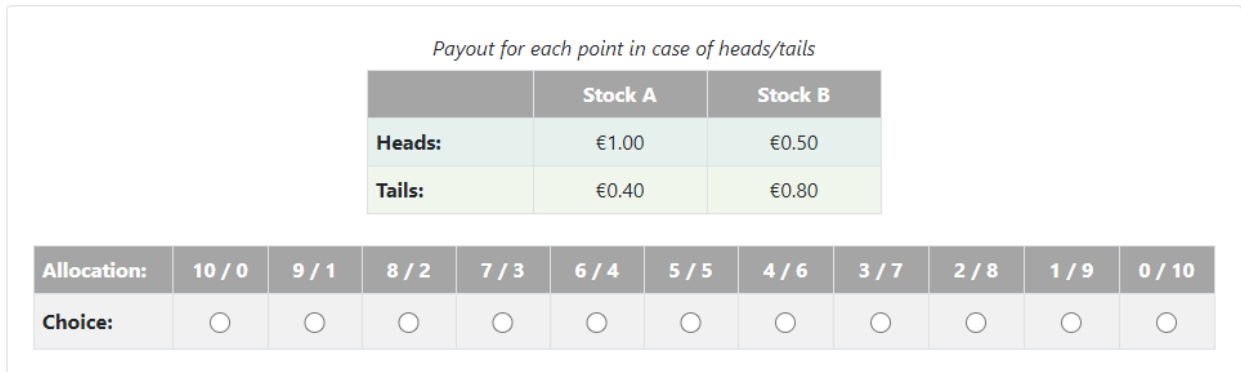
Figure A1: Screenshot of the decision environment in treatment AT.

Figure A2: Screenshot of the decision environment in treatment AG.

Allocation:	10 / 0	9 / 1	8 / 2	7 / 3	6 / 4	5 / 5	4 / 6	3 / 7	2 / 8	1 / 9	0 / 10
					Payout for each point in case of heads/tails						
Heads:	€1.00	€0.95	€0.90	€0.85	€0.80	€0.75	€0.70	€0.65	€0.60	€0.55	€0.50
Tails:	€0.40	€0.44	€0.48	€0.52	€0.56	€0.60	€0.64	€0.68	€0.72	€0.76	€0.80
Choice:	○	○	○	○	○	○	○	○	○	○	○

Figure A3: Screenshot of the decision environment in treatment PT.
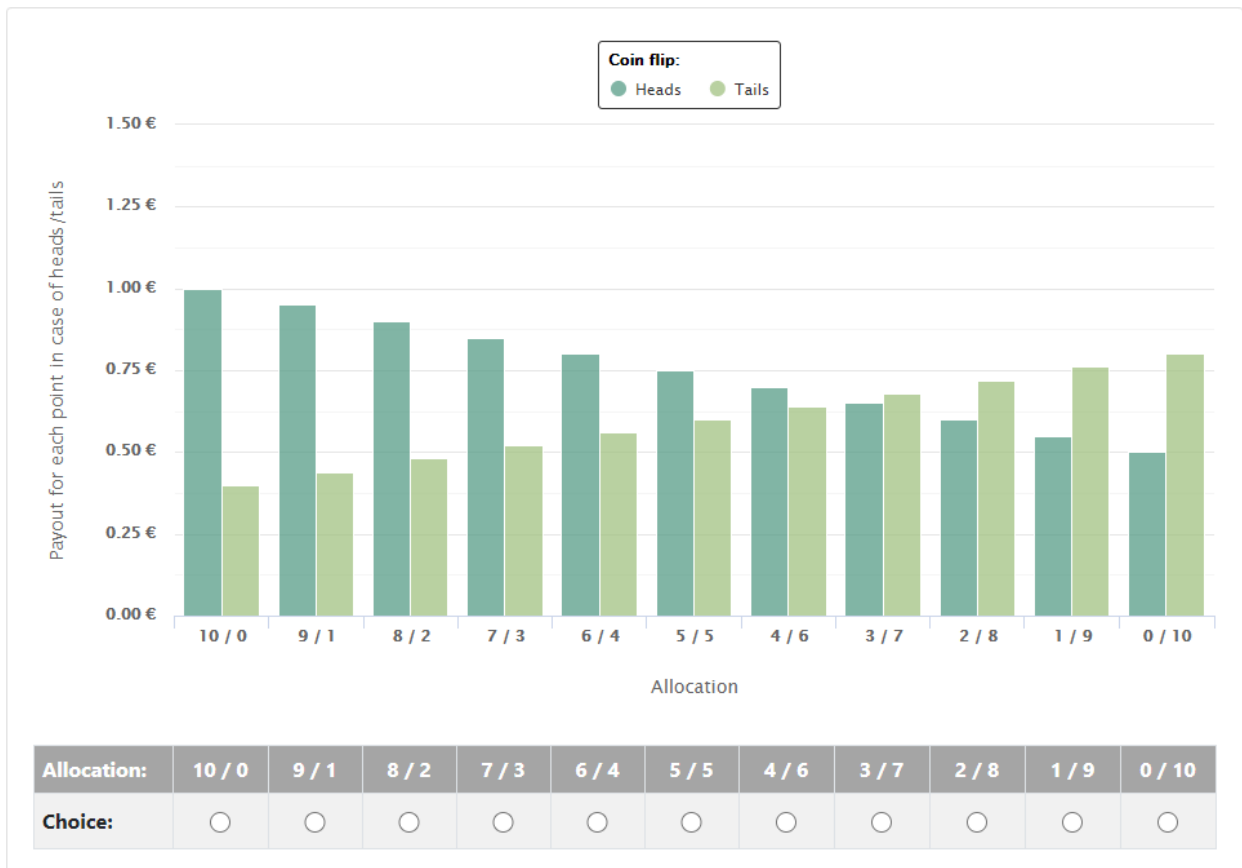
Figure A4: Screenshot of the decision environment in treatment PG.

B. Decision-Making Quality Measures

In each opportunity set $j \in \{1, 2, \ldots, 14\}$ (i.e., in each decision), participant i is endowed with €10 to be allocated to assets $s \in \{A, B\}$. Let $n_{i,j,s}$ denote the euro amount allocated to asset s. Participants have to invest their entire endowment, hence $\sum_s n_{i,j,s} = 10$.

The return per € invested in asset s if the coin comes up heads is denoted as $h_{j,s}$; the return per € invested if it comes up tails is denoted as $t_{j,s}$. Thus, the return (in €) of participant i's allocation in opportunity set j will either be

$$h_{i,j} = \sum_s n_{i,j,s} \cdot h_{j,s} \qquad \text{if the coin comes up heads, or}$$

$$t_{i,j} = \sum_s n_{i,j,s} \cdot t_{j,s} \qquad \text{if the coin comes up tails.}$$

Let the tuple $\boldsymbol{x}_{i,j} = (h_{i,j}, t_{i,j})$ denote the payoff profile of participant i's asset allocation in opportunity set j. Since in our experiment the relation between asset allocation and portfolio payoff is bijective, the payoff profile $\boldsymbol{x}_{i,j}$ also uniquely identifies the participant's asset allocation. We hence use the term portfolio (choice) interchangeably for the participant's payoff profile (as defined above) and their asset allocation.

We define three measures of decision-making quality: (i) violations of first-order stochastic dominance (*FOSD*), (ii) violations of state-by-state dominance (*SSD*), and (iii) violations of the generalized axiom of revealed preferences (*GARP*). The measures used in our experiment are inspired by Holzmeister et al. (2019), but we expand by introducing state-by-state dominance and by offering concise formulae for all measures, which hold for both continuous and discrete cases.

Expected return. The expected return of participant i's portfolio choice in opportunity set j is given by $ER_{i,j} = \frac{(h_{i,j} + t_{i,j})}{2}$. We define participant i's mean expected return, ER_i, as the average across all opportunity sets, $ER_i = \frac{1}{14} \cdot \sum_{j=1}^{14} ER_{i,j}$. Finally, in the main text, we denote by ER the set of mean expected portfolio returns of all participants.

Standard deviation. As a measure of portfolio risk, we calculate the standard deviation of participant i's portfolio in opportunity set j. As the two possible outcomes $h_{i,j}$ and $t_{i,j}$ both have probability $\frac{1}{2}$, the standard deviation is:

$$SD_{i,j} = \sqrt{\frac{h_{i,j}^2 + t_{i,j}^2}{2} - \left(\frac{h_{i,j} + t_{i,j}}{2}\right)^2} = \frac{|h_{i,j} - t_{i,j}|}{2}.$$

We define participant i's average portfolio risk, SD_i, as the average across all opportunity sets, $SD_i = \frac{1}{14} \cdot \sum_{j=1}^{14} SD_{i,j}$. As above, SD denotes the set of mean portfolio risks of all participants.

Coefficient of variation. As a measure of portfolio efficiency, we calculate the coefficient of variation of participant i's portfolio allocation in opportunity set j, the ratio of $SD_{i,j}$ to $ER_{i,j}$: $CV_{i,j} = \frac{SD_{i,j}}{ER_{i,j}}$. $CV_{i,j}$ measures the unit of risk taken per unit of expected return in opportunity set j and corresponds to the inverse of the portfolio's risk-adjusted return. Thus, comparing two allocations with the same expected return *or* the same portfolio risk, a higher (lower) value of $CV_{i,j}$ implies a less (more) efficient portfolio. The reason for using the coefficient of variation rather than return per unit of risk is that some opportunity sets allow for perfectly hedged allocations (i.e., a standard deviation of zero), which would give rise to undefined risk-adjusted returns.

We define participant i's average coefficient of variation, CV_i, as the average across all opportunity sets, $CV_i = \frac{1}{14} \cdot \sum_{j=1}^{14} CV_{i,j}$. CV denotes the set of mean coefficients of variation across of all participants.

Violations of first-order stochastic dominance. Suppose participant i chooses portfolio $\boldsymbol{x}_{i,j} = (h_{i,j}, t_{i,j})$ in opportunity set j, i.e., an allocation returning $h_{i,j}$ or $t_{i,j}$ with 50% probability each, as illustrated in Figure B1. For the given opportunity set, the figure also depicts an alternative portfolio $\boldsymbol{x}'_{i,j} = (h'_{i,j}, t'_{i,j})$. Apparently, portfolio $\boldsymbol{x}'_{i,j}$ first-order stochastically dominates $\boldsymbol{x}_{i,j}$ because $t'_{i,j} = h_{i,j}$ and $h'_{i,j} > t_{i,j}$.
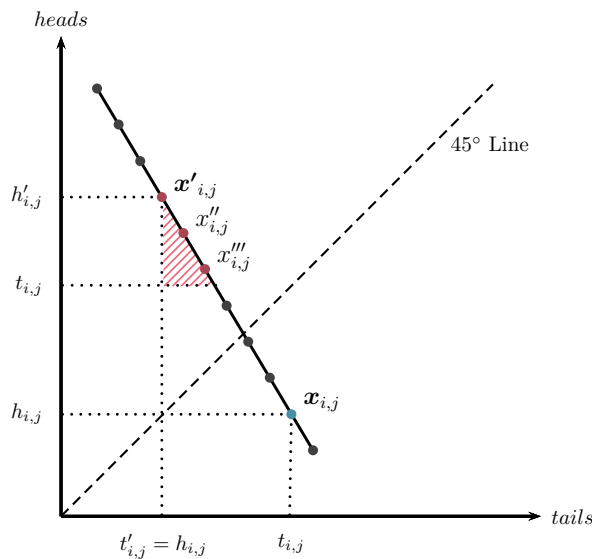
Figure B1: Violation of the principle of first-order stochastic dominance. The chosen portfolio $\boldsymbol{x}_{i,j}$ is first-order stochastically dominated by the alternative portfolio $\boldsymbol{x}'_{i,j}$ because $t'_{i,j} = h_{i,j}$ and $h'_{i,j} > t_{i,j}$. That is, portfolio $\boldsymbol{x}'_{i,j}$ yields a higher expected return than $\boldsymbol{x}_{i,j}$ while guaranteeing the same minimum payoff. The shaded area shows the dominating alternatives if the choice set was continuous. The corresponding dots on the intersection of the shaded area and the portfolio line constitute the set of feasible portfolios dominating the chosen portfolio $\boldsymbol{x}_{i,j}$. $\boldsymbol{x}_{i,j}$ is also first-order stochastically dominated by $\boldsymbol{x}''_{i,j}$ and $\boldsymbol{x}'''_{i,j}$, but $\boldsymbol{x}'_{i,j}$ maximizes the expected payoff, as is apparent from the slope of the feasible set.

Following Banks et al. (2018), we use the difference between the maximum expected return of a portfolio that provides the same minimum payoff as the chosen portfolio and the expected return of

the chosen portfolio as a measure of how closely participant i's choice in opportunity set j complies with the principle of *FOSD* (Hadar and Russell, 1969), i.e., we define:

$$FOSD_{i,j} = \max_{h'_{i,j}, t'_{i,j}} \left(\frac{h'_{i,j} + t'_{i,j}}{2} - \frac{h_{i,j} + t_{i,j}}{2} \right), \quad \text{s.t. } min(h'_{i,j}, t'_{i,j}) \geq min(h_{i,j}, t_{i,j}).$$

Hence, in Figure B1, portfolio $\boldsymbol{x}'_{i,j}$ determines the value of the *FOSD* measure for a participant choosing portfolio $\boldsymbol{x}_{i,j}$. In contrast, the concept of state-by-state dominance (*SSD*) has the same objective function, but rests on a stricter optimization constraint:

$$SSD_{i,j} = \max_{h'_{i,j}, t'_{i,j}} \left(\frac{h'_{i,j} + t'_{i,j}}{2} - \frac{h_{i,j} + t_{i,j}}{2} \right), \quad \text{s.t. } h'_{i,j} \geq h_{i,j} \text{ and } t'_{i,j} \geq t_{i,j}.$$

Note that the above formalization captures (i) the property that state-by-state dominance is a subset of first-order stochastic dominance and (ii) the property that both *FOSD* and *SSD* are bounded from below by 0 (in case the portfolio $\boldsymbol{x}_{i,j} = (h_{i,j}, t_{i,j})$ is not *FOSD* or *SSD* dominated, respectively, by any other portfolio). In the two-state two-asset case, a *portfolio* state-by-state violation ($SSD_{i,j} > 0$) is only possible if the slope of the feasible set is positive, i.e., if one of the *assets* is state-by-state dominated by the other. Consequently, *FOSD* and *SSD* are distinct with regards to which allocation decisions they are applied to in our experiment. *FOSD* is calculated for all allocation decisions showing negatively sloped opportunity sets. *SSD* is calculated for the remaining allocation decisions showing positively sloped opportunity sets.

For each participant, we compute the corresponding averages as $FOSD_i = \frac{1}{14} \cdot \sum_{j=1}^{14} FOSD_{i,j}$ and $SSD_i = \frac{1}{14} \cdot \sum_{j=1}^{14} SSD_{i,j}$. *FOSD* and *SSD* denote the set of mean first-order stochastic dominance and state-by-state dominance violations, respectively, of all participants.

Violations of the generalized axiom of revealed preference. In the following, we transfer the implications of the generalized axiom of revealed preference (*GARP*) to the investment task in the present study.[9] *GARP* requires that participant i facing two alternatives $\boldsymbol{x}_{i,j}$ and $\boldsymbol{x}'_{i,j}$ which are feasible in opportunity set j, cannot both directly revealed prefer $\boldsymbol{x}_{i,j}$ over $\boldsymbol{x}'_{i,j}$ and strictly and directly revealed prefer $\boldsymbol{x}'_{i,j}$ over $\boldsymbol{x}_{i,j}$ in any other opportunity set k that contains both alternatives. In the continuous textbook case, *GARP* violations are typically illustrated by inconsistent choices under two different price regimes for two goods (which, in line with prior notation in this paper, we will call h and t). Monotonic preferences mandate that optimal bundles lie on the budget line. If shifts in relative prices lead to intersecting budget lines, parts of either budget line will lie in the interior of the other budget set. To develop the discrete case, suppose first that Figure B2 refers to a continuous decision space of consumption bundles and that participant i chooses $\boldsymbol{x}_{i,j}$ in opportunity set j as defined by the budget line b_j. Thereby, i reveals to weakly prefer alternative/bundle $\boldsymbol{x}_{i,j}$ over all other alternatives on b_j. Importantly, she hence also weakly prefers $\boldsymbol{x}_{i,j}$ over all alternatives

---

[9] Banerjee and Murphy (2009) show that for the two-commodity case, the weak generalized axiom of revealed preference (*WGARP*) is equivalent to the generalized axiom of revealed preference (*GARP*). Technically, be build our arguments on the axiom of *WGARP*.

on b_j *northeast of* $\boldsymbol{x}'_{i,j}$, which, by monotonicity, are strictly preferred over $\boldsymbol{x}'_{i,j}$. So, $\boldsymbol{x}_{i,j}$ must be strictly preferred over $\boldsymbol{x}'_{i,j}$. In (the continuous) opportunity set k, spanned by budget line b_k, it is apparent that both alternatives $\boldsymbol{x}_{i,j}$ and $\boldsymbol{x}'_{i,j}$ are available to the decision maker. If participant i now chooses $\boldsymbol{x}'_{i,j}$, she reveals to weakly prefer this alternative over $\boldsymbol{x}_{i,j}$ (in the chosen illustration, as above, the extension to strict preference can be made), which is *GARP*-violating her revealing to strictly prefer $\boldsymbol{x}_{i,j}$ over $\boldsymbol{x}'_{i,j}$ in opportunity set j.
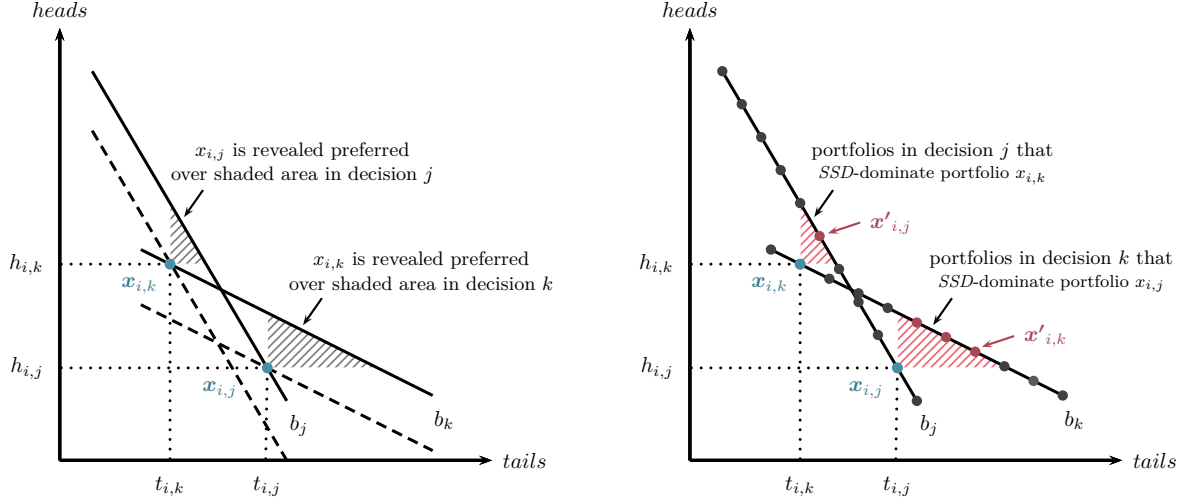
Figure B2: Violation of the generalized axiom of revealed preferences in the continuous two-commodity case (left panel) and the discrete investment task (right panel). The two panels outline the similarities and differences between *MPI* and *GARP*. A decision maker choosing $\boldsymbol{x}_{i,j}$ in opportunity set j and $\boldsymbol{x}_{i,k}$ in opportunity set k violates the generalized axiom of revealed preferences. In the continuous commodity case, *MPI* is computed as the wealth the decision maker loses to an arbitrageur who promises to deliver the same bundles, but buys them under the opposite price regimes (illustrated by the dashed budget lines). In the discrete investment task, to compute *GARP*, we need to adapt the concept as portfolio $\boldsymbol{x}_{i,j}$ is not available in decision k and, vice versa, $\boldsymbol{x}_{i,k}$ is not available in decision j. Here, the arbitrageur buys the right to make the portfolio decisions by promising to deliver the payoffs the decision maker would obtain based on the respective choices. Then, the arbitrageur maximizes expected payoff by choosing $\boldsymbol{x}'_{i,j}$ in decision j (which *SSD*-dominates $\boldsymbol{x}_{i,k}$) and $\boldsymbol{x}'_{i,k}$ in decision k (which *SSD*-dominates $\boldsymbol{x}_{i,j}$). We define the *GARP* measure as the difference in expected payoffs between the arbitrageur and the decision maker.

Recalling that the two opportunity sets j and k are defined by two price vectors $\boldsymbol{p}_j$ and $\boldsymbol{p}_k$ with $p_j^h < p_k^h$ and $p_j^t > p_k^t$, Echenique et al. (2011) propose to measure the extent of violations of *GARP* by the Money Pump Index (*MPI*), which is based on the idea that an arbitrageur can exploit violations in revealed preferences: The arbitrageur could make profit by buying bundle $\boldsymbol{x}_{i,j}$ as part of opportunity set j at prices $\boldsymbol{p}_j$ and then selling it as part of opportunity set k at prices $\boldsymbol{p}_k$; likewise, the arbitrageur could buy bundle $\boldsymbol{x}'_{i,j}$ as part of opportunity set k at prices p_k and sell it as part of opportunity set j at prices p_j. The resulting profit is what they define as the Money Pump Index.

$$GARP_{i,j,k} = h_{i,j}(p_k^h - p_j^h) + t_{i,j}(p_k^t - p_j^t) + h'_{i,j}(p_j^h - p_k^h) + t'_{i,j}(p_j^t - p_k^t)$$
$$= (h'_{i,j} - h_{i,j})(p_k^h - p_j^h) + (t'_{i,j} - t_{i,j})(p_k^t - p_j^t)$$

which is always positive, as it is the sum of the products of two positives and two negatives. Note that this assumes the arbitrageur can buy at $\boldsymbol{p}_j$ and sell at $\boldsymbol{p}_k$, so he has access to both prices at the same time.

Transferring this concept from the continuous decision space of consumption bundles (under two price regimes) to our experiment where participants invest €10 in (two times) two 50/50-lotteries, the budget sets of possible consumption bundles of two goods translate to discrete sets of combinations of payments the participant can obtain in the two states of nature *heads* (h) and *tails* (t). While the principle idea to price *GARP* violations remains the same, some adjustments need to be made. As can be seen from Figure B2, due to the discreteness of the choice sets, their intersection is empty. I.e., $\boldsymbol{x}_{i,j}$ is in fact not available in opportunity set k. A (risk neutral) arbitrageur must ensure he is able to pay off the participant regardless of the state of the world. Technically, in our case, the arbitrageur will *buy the right to make the portfolio choice* from the participant by promising to pay her according to her choice. His profit is generated by finding the expected-value-maximizing portfolio that allows him to fulfill his promise. As implicitly required in the derivation of *MPI*, we will also assume that the arbitrageur has access to both price regimes (i.e., here, both two lottery pairs) at the same time. In other words, he can buy the participant's right to make a portfolio choice in opportunity set j and use the outcome of the lottery pair in k to pay her off (and, *vice versa*, buy the participant's decision right in k, using the outcome of j to pay her off). Consequently, the arbitrageur must ensure sufficient liquidity in either portfolio choice. Note that the simultaneity of access to prices/lottery pairs *does not* mean that the outcomes are aggregated, keeping the concept in the spirit of Echenique et al. (2011).

Formally, an arbitrageur will thus maximize the following "crosswise" objective function:

$$GARP_{i,j,k} = \max_{h'_{i,k}, t'_{i,k}, h'_{i,j}, t'_{i,j}} \left[\left(\frac{h'_{i,k} + t'_{i,k}}{2} - \frac{h_{i,j} + t_{i,j}}{2} \right) + \left(\frac{h'_{i,j} + t'_{i,j}}{2} - \frac{h_{i,k} + t_{i,k}}{2} \right) \right],$$

$$\text{s.t. } h'_{i,k} \geq h_{i,j}, t'_{i,k} \geq t_{i,j}, h'_{i,j} \geq h_{i,k}, \text{ and } t'_{i,j} \geq t_{i,k}.$$

In essence, the arbitrage-argument of *GARP*-violations can thus also be described as *SSD*-violations *across choices*.

We calculate the *GARP*-violation index for $\binom{14}{2} = 91$ pairwise combinations of opportunity sets. We define participant i's average *GARP* index, $GARP_i$, as the average across all pairwise combinations, $GARP_i = \frac{1}{91} \cdot \sum_{j=2}^{14} \sum_{k=1}^{j-1} GARP_{i,j,k}$. *GARP* denotes the set of mean revealed preference violation indices of all participants.

C. Supplementary Tables

Below, we present the regression estimates underlying the graphical illustration of our results presented in the main text, as well as robustness checks and supplementary analyses as referenced in the main text. In particular, Table C1 shows the estimates of ordinary least squares regressions of the three decision-making quality indices (*FOSD*, *SSD*, and *GARP*) on the treatment indicators P and G and their interaction term, corresponding to Figure 2 in the main text. As a robustness check for the results reported in Table C1, Table C2 summarizes the results of random effects regressions on the non-aggregated level, i.e., on a per-decision basis.

Similarly, Table C3 depicts the estimates of ordinary least squares regressions of the three performance measures (*ER*, *SD*, and *CV*) on the treatment indicators P and G and their interaction term, corresponding to Figure 3 in the main text. Again, as a robustness check for the results reported in Table C3, Table C4 shows the results of random effects regressions on a per-decision basis.

Table C5 (Table C6) shows the results of ordinary least squares estimates of the three decision-making quality indices (three performance measures) on the the treatment indicators P and G and their interaction term, as well as the two cognitive measures *APM* and *NUM* and their interaction terms with the treatment variables. The estimates in Table C5 (Table C6) form the basis for the graphical representation of results in both Figures 4 and 5 (both Figures 6 and 7) in the main text.

Table C7 shows the Pearson correlation coefficients between the six outcome variables analyzed in the main text (*FOSD*, *SSD*, *GARP*, *ER*, *SD*, and *CV*) and participants' survey responses on perceived risk, perceived difficulty, and satisfaction. Table C8 shows the results of ordinary least squares regressions of perceived risk, perceived difficulty, and satisfaction on the the treatment indicators P and G and their interaction term.

Table C1: Ordinary least squares regressions of **(i)** violations of first-order stochastic dominance (*FOSD*), **(ii)** violations of state-by-state dominance (*SSD*), and **(iii)** violations of the generalized axiom of revealed preferences on indicators for the treatments "Portfolio" (*P*) and "Graphical" (*G*), and their interaction term (*P*×*G*). The three dependent variables are *z*-standardized scores. Robust standard errors are provided in parentheses. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

	FOSD	SSD	GARP
Treatment Indicators:			
P ("Portfolio")	−0.111	−0.083	−0.008
	(0.121)	(0.154)	(0.135)
G ("Graphical")	0.349	0.306	0.431**
	(0.186)	(0.162)	(0.183)
Interaction Effect:			
P×G	−0.502*	−0.422	−0.527*
	(0.218)	(0.219)	(0.220)
Constant	0.014	0.000	−0.077
	(0.091)	(0.109)	(0.102)
Linear Combinations:			
P + P×G	−0.613***	−0.505**	−0.536**
	(0.181)	(0.156)	(0.173)
G + P×G	−0.153	−0.117	−0.096
	(0.113)	(0.148)	(0.121)
Observations	325	325	325
Adj. R^2	0.044	0.027	0.035
$F(3, 321)$	4.345**	3.660*	3.211*

Table C2: Random effects regressions of **(i)** violations of first-order stochastic dominance (*FOSD*; 11 decisions per participant) and **(ii)** violations of state-by-state dominance (*SSD*; 3 decisions per participant) on indicators for the treatments "Portfolio" (*P*) and "Graphical" (*G*), and their interaction term (*P*×*G*). The two dependent variables are *z*-standardized scores. Robust standard errors, clustered on the participant level ($n = 325$ clusters), are provided in parentheses. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

	FOSD	SSD
Treatment Indicators:		
P ("Portfolio")	−0.037	−0.063
	(0.040)	(0.117)
G ("Graphical")	0.117	0.233
	(0.062)	(0.123)
Interaction Effect:		
P×G	−0.168*	−0.322
	(0.073)	(0.166)
Constant	0.005	0.000
	(0.031)	(0.082)
Linear Combinations:		
P + P×G	−0.206***	−0.385**
	(0.061)	(0.118)
G + P×G	−0.051	−0.089
	(0.038)	(0.112)
Observations	3,575	975
Adj. R^2	0.005	0.018
$F(3, 321)$	4.382**	3.683*

Table C3: Ordinary least squares regressions of **(i)** the mean expected portfolio return (*ER*), **(ii)** the mean portfolio standard deviation (*SD*), and **(iii)** the mean portfolio variation coefficient (*CV*) on indicators for the treatments "Portfolio" (*P*) and "Graphical" (*G*), and their interaction term ($P \times G$). The three dependent variables are z-standardized scores. Robust standard errors are provided in parentheses. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

	ER	SD	CV
Treatment Indicators:			
P ("Portfolio")	−0.087	−0.355*	−0.386*
	(0.149)	(0.154)	(0.152)
G ("Graphical")	−0.280	0.161	0.230
	(0.165)	(0.136)	(0.134)
Interaction Effect:			
$P \times G$	0.403	−0.147	−0.242
	(0.221)	(0.215)	(0.212)
Constant	0.081	0.144	0.150
	(0.103)	(0.100)	(0.097)
Linear Combinations:			
$P + P \times G$	0.316	−0.502***	−0.628***
	(0.163)	(0.149)	(0.147)
$G + P \times G$	0.124	0.013	−0.012
	(0.147)	(0.166)	(0.164)
Observations	325	325	325
Adj. R^2	0.006	0.041	0.064
$F(3, 321)$	1.409	5.976***	8.799***

Table C4: Random effects regressions of **(i)** the mean expected portfolio return (*ER*), **(ii)** the mean portfolio standard deviation (*SD*), and **(iii)** the mean portfolio variation coefficient (*CV*) on indicators for the treatments "Portfolio" (*P*) and "Graphical" (*G*), and their interaction term ($P \times G$). The three dependent variables are z-standardized scores. Robust standard errors, clustered on the participant level ($n = 325$ clusters), are provided in parentheses. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

	ER	SD	CV
Treatment Indicators:			
P ("Portfolio")	−0.039	−0.157*	−0.168*
	(0.066)	(0.068)	(0.066)
G ("Graphical")	−0.125	0.071	0.100
	(0.074)	(0.060)	(0.058)
Interaction Effect:			
$P \times G$	0.181	−0.065	−0.105
	(0.099)	(0.094)	(0.092)
Constant	0.036	0.064	0.065
	(0.046)	(0.044)	(0.042)
Linear Combinations:			
$P + P \times G$	0.142	−0.222***	−0.273***
	(0.073)	(0.066)	(0.064)
$G + P \times G$	0.055	0.006	−0.005
	(0.065)	(0.073)	(0.071)
Observations	4,550	4,550	4,550
Adj. R^2	0.002	0.009	0.013
$F(3, 321)$	1.421	6.027***	8.875***

Table C5: Ordinary least squares regressions of **(i)** violations of first-order stochastic dominance (*FOSD*), **(ii)** violations of state-by-state dominance (*SSD*), and **(iii)** violations of the generalized axiom of revealed preferences (*GARP*) on indicators for the treatments "Portfolio" (*P*) and "Graphical" (*G*), participants' numeracy skills (*NUM*) and fluid intelligence (*APM*), and the interaction terms of treatment indicators and cognitive measures. The three dependent variables are z-standardized scores. Robust standard errors are provided in parentheses. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

	FOSD	*SSD*	*GARP*
Treatment Indicators:			
P ("Portfolio")	−0.802	−0.377	−0.557
	(0.502)	(0.573)	(0.514)
G ("Graphical")	0.794	0.026	0.198
	(0.681)	(0.583)	(0.623)
Cognitive Measures:			
APM (Fluid Intelligence)	−0.143	−0.060	−0.105
	(0.082)	(0.072)	(0.080)
NUM (Numeracy)	−0.135**	−0.242***	−0.149**
	(0.047)	(0.070)	(0.051)
Interaction Effects:			
P×G	−0.868	−0.307	−0.104
	(0.856)	(0.837)	(0.802)
P×APM	0.052	−0.132	−0.036
	(0.091)	(0.096)	(0.094)
P×NUM	0.083	0.189	0.139
	(0.070)	(0.098)	(0.077)
G×APM	−0.062	−0.045	−0.047
	(0.138)	(0.105)	(0.131)
G×NUM	−0.024	0.109	0.094
	(0.143)	(0.100)	(0.123)
P×G×APM	0.052	0.125	0.099
	(0.154)	(0.134)	(0.147)
P×G×NUM	0.026	−0.153	0.220
	(0.161)	(0.141)	(0.150)
Constant	1.866**	1.314*	1.816**
	(0.601)	(0.538)	(0.659)
RT Controls	*yes*	*yes*	*yes*
Observations	325	325	325
Adj. R^2	0.144	0.157	0.129
$F(17, 307)$	4.425***	5.284***	5.410***

To be continued on next page.

Continued from Table C5. Linear combinations (based on the regression estimates provided in Table C5) isolating the effects of *APM* and *NUM* for each of the four treatment conditions and the differences between effects for all relevant treatment comparisons.

Effects Separated by Treatments:			
APM in *AT*	−0.143	−0.060	−0.105
	(0.082)	(0.072)	(0.080)
APM in *PT*	−0.892	−0.568	−0.698
	(0.499)	(0.553)	(0.505)
APM in *AG*	0.589	−0.079	0.046
	(0.670)	(0.557)	(0.600)
APM in *PG*	−0.108	−0.463	−0.447
	(1.036)	(0.998)	(1.015)
NUM in *AT*	−0.135**	−0.242***	−0.149**
	(0.047)	(0.070)	(0.051)
NUM in *PT*	−0.853	−0.430	−0.567
	(0.487)	(0.549)	(0.500)
NUM in *AG*	0.636	−0.107	0.144
	(0.640)	(0.561)	(0.599)
NUM in *PG*	−0.057	−0.448	−0.494
	(1.036)	(0.997)	(1.012)
Differences in Effects:			
APM in *AT* vs. *PT*	−0.750	−0.509	−0.593
	(0.443)	(0.53)	(0.458)
APM in *AT* vs. *AG*	0.732	−0.019	0.152
	(0.632)	(0.537)	(0.563)
APM in *PT* vs. *PG*	0.784	0.106	0.251
	(0.684)	(0.589)	(0.626)
APM in *AG* vs. *PG*	−0.698	−0.384	−0.494
	(0.522)	(0.580)	(0.532)
NUM in *AT* vs. *PT*	−0.719	−0.188	−0.418
	(0.483)	(0.529)	(0.491)
NUM in *AT* vs. *AG*	0.770	0.135	0.293
	(0.635)	(0.539)	(0.589)
NUM in *PT* vs. *PG*	0.796	−0.018	0.073
	(0.686)	(0.592)	(0.628)
NUM in *AG* vs. *PG*	−0.692	−0.341	−0.638
	(0.518)	(0.583)	(0.525)

Table C6: Ordinary least squares regressions of **(i)** the mean expected portfolio return (ER), **(ii)** the mean portfolio standard deviation (SD), and **(iii)** the mean portfolio variation coefficient (CV) on indicators for the treatments "Portfolio" (P) and "Graphical" (G), participants' numeracy skills (NUM) and fluid intelligence (APM), and the interaction terms of treatment indicators and cognitive measures. The three dependent variables are z-standardized scores. Robust standard errors are provided in parentheses. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

	ER	SD	CV
Treatment Indicators:			
P ("Portfolio")	0.371	−0.990	−1.183*
	(0.464)	(0.548)	(0.568)
G ("Graphical")	0.048	−0.252	−0.296
	(0.532)	(0.563)	(0.570)
Cognitive Measures:			
APM (Fluid Intelligence)	0.090	−0.080	−0.110
	(0.056)	(0.071)	(0.074)
NUM (Numeracy)	0.223***	0.068	0.032
	(0.047)	(0.037)	(0.039)
Interaction Effects:			
$P{\times}G$	−0.096	0.170	0.226
	(0.766)	(0.838)	(0.839)
$P{\times}APM$	0.077	0.059	0.051
	(0.082)	(0.099)	(0.101)
$P{\times}NUM$	−0.164*	0.069	0.107
	(0.080)	(0.082)	(0.084)
$G{\times}APM$	0.036	0.127	0.130
	(0.096)	(0.093)	(0.096)
$G{\times}NUM$	−0.110	−0.058	−0.039
	(0.097)	(0.071)	(0.072)
$P{\times}G{\times}APM$	−0.097	−0.091	−0.081
	(0.127)	(0.135)	(0.136)
$P{\times}G{\times}NUM$	0.203	0.049	0.011
	(0.136)	(0.120)	(0.121)
Constant	−1.706**	−0.126	0.258
	(0.516)	(0.455)	(0.494)
RT Controls	*yes*	*yes*	*yes*
Observations	325	325	325
Adj. R^2	0.164	0.120	0.132
$F(17, 307)$	6.497***	5.115***	5.352***

Continued from Table C6. Linear combinations (based on the regression estimates provided in Table C6) isolating the effects of *APM* and *NUM* for each of the four treatment conditions and the differences between effects for all relevant treatment comparisons.

Effects Separated by Treatments:			
APM in *AT*	0.090	−0.080	−0.110
	(0.056)	(0.071)	(0.074)
APM in *PT*	0.538	−1.011	−1.241*
	(0.45)	(0.525)	(0.545)
APM in *AG*	0.174	−0.205	−0.275
	(0.514)	(0.541)	(0.549)
APM in *PG*	0.525	−1.227	−1.488
	(0.869)	(0.980)	(1.008)
NUM in *AT*	0.223***	0.068	0.032
	(0.047)	(0.037)	(0.039)
NUM in *PT*	0.430	−0.853	−1.044
	(0.449)	(0.530)	(0.549)
NUM in *AG*	0.161	−0.241	−0.303
	(0.503)	(0.541)	(0.550)
NUM in *PG*	0.571	−1.113	−1.368
	(0.867)	(0.978)	(1.006)
Differences in Effects:			
APM in *AT* vs. *PT*	0.448	−0.932	−1.132*
	(0.423)	(0.480)	(0.498)
APM in *AT* vs. *AG*	0.084	−0.125	−0.165
	(0.491)	(0.497)	(0.502)
APM in *PT* vs. *PG*	−0.013	−0.216	−0.247
	(0.538)	(0.573)	(0.579)
APM in *AG* vs. *PG*	0.351	−1.023	−1.213*
	(0.477)	(0.556)	(0.575)
NUM in *AT* vs. *PT*	0.207	−0.921	−1.076*
	(0.433)	(0.523)	(0.543)
NUM in *AT* vs. *AG*	−0.063	−0.309	−0.334
	(0.487)	(0.533)	(0.542)
NUM in *PT* vs. *PG*	0.141	−0.260	−0.323
	(0.539)	(0.572)	(0.579)
NUM in *AG* vs. *PG*	0.410	−0.872	−1.065
	(0.479)	(0.559)	(0.578)

Table C7: Pearson correlation coefficients between (i) violations of first-order stochastic dominance (*FOSD*), (ii) violations of state-by-state dominance (*SSD*), (iii) violations of the generalized axiom of revealed preferences (*GARP*), (iv) the mean expected portfolio return (*ER*), (v) the mean portfolio standard deviation (*SD*), and (vi) the mean portfolio variation coefficient (*CV*) presented in rows and (a) perceived riskiness, (b) perceived difficulty, and (c) satisfaction presented in columns. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

	Perc. Risk	*Perc. Diff.*	*Satisfaction*
FOSD	0.110*	0.038	−0.171**
SSD	0.030	0.170**	−0.271***
GARP	0.072	0.076	−0.212***
ER	0.150**	−0.170**	0.275***
SD	0.587***	−0.116*	0.118*
CV	0.605***	−0.084	0.063

Table C8: Ordinary least squares regressions of **(i)** perceived riskiness, **(ii)** perceived difficulty, and **(iii)** satisfaction on indicators for the treatments "Portfolio" (*P*) and "Graphical" (*G*), and their interaction term (*P×G*). The three dependent variables are *z*-standardized scores. Robust standard errors are provided in parentheses. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

	Perc. Risk	*Perc. Diff.*	*Satisfaction*
Treatment Indicators:			
P ("Portfolio")	−0.274	−0.074	0.109
	(0.157)	(0.157)	(0.175)
G ("Graphical")	0.268	0.103	−0.150
	(0.154)	(0.162)	(0.172)
Interaction Effect:			
P×G	0.031	0.213	−0.022
	(0.218)	(0.222)	(0.226)
Constant	−0.004	−0.074	0.027
	(0.108)	(0.112)	(0.133)
Linear Combinations:			
P + P×G	−0.244	0.139	0.087
	(0.151)	(0.156)	(0.143)
G + P×G	0.299	0.316*	−0.172
	(0.155)	(0.151)	(0.146)
Observations	325	325	325
Adj. R^2	0.028	0.006	0.000
$F(3, 321)$	3.926**	1.655	0.962

2021-15 **Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan:** The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

2021-14 **Jeppe Christoffersen, Felix Holzmeister, Thomas Plenborg:** What is Risk to Managers?

2021-13 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Trust in health care credence goods: Experimental evidence on framing andsubject pool effects

2021-12 **Rene Schwaiger, Laura Hueber:** Do MTurkers Exhibit Myopic Loss Aversion?

2021-11 **Felix Holzmeister, Christoph Huber, Stefan Palan:** A Critical Perspective on the Conceptualization of Risk in Behavioral and Experimental Finance

2021-10 **Michael Razen, Alexander Kupfer:** Can increased tax transparency curb corporate tax avoidance?

2021-09 **Changxia Ke, Florian Morath, Anthony Newell, Lionel Page:** Too big to prevail: The paradox of power in coalition formation

2021-08 **Marco Haan, Pim Heijnen, Martin Obradovits:** Competition with List Prices

2021-07 **Martin Dufwenberg, Olof Johansson-Stenman, Michael Kirchler, Florian Lindner, Rene Schwaiger:** Mean Markets or Kind Commerce?

2021-06 **Christoph Huber, Jürgen Huber, and Michael Kirchler:** Volatility Shocks and Investment Behavior

2021-05 **Max Breitenlechner, Georgios Georgiadis, Ben Schumann:** What goes around comes around: How large are spillbacks from US monetary policy?

2021-04 **Utz Weitzel, Michael Kirchler:** The Banker's Oath And Financial Advice

2021-03 **Martin Holmen, Felix Holzmeister, Michael Kirchler, Matthias Stefan, Erik Wengström:** Economic Preferences and Personality Traits Among Finance Professionals and the General Population

2021-02 **Christian König-Kersting:** On the Robustness of Social Norm Elicitation

2021-01 **Laura Hueber, Rene Schwaiger:** Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

University of Innsbruck

Working Papers in Economics and Statistics

2021-15

Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan

The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

Abstract
The question of optimal presentation format and choice architecture for investment decisions has gained momentum among researchers, policy makers, and practitioners alike. Motivated by the question how to provide information to investors in a way to improve financial decision-making, we conduct an investment experiment. We implement a 2 x 2 factorial design to test the effect of presentation format (graphical vs. tabular) and choice architecture/complexity (asset selection vs. predefined portfolios) on decision-making quality. Overall, our results suggest a differential effect of presentation format and choice architecture: Firstly, we find that the graphical presentation format lowers decision-making quality when the environment is more complex (asset selection). Secondly, within graphical presentation, decision quality is higher when the choice architecture is simplified (predefined portfolios). In addition, we find that a simplified choice architecture leads to higher risk-adjusted returns and reduces the relevance of fluid intelligence and numeracy for decision-making quality.