

Farago, Adam; Holmen, Martin; Holzmeister, Felix; Kirchler, Michael; Razen, Michael

Working Paper

Cognitive skills and economic preferences in the fund industry

Working Papers in Economics and Statistics, No. 2019-16

Provided in Cooperation with:

Institute of Public Finance, University of Innsbruck

Suggested Citation: Farago, Adam; Holmen, Martin; Holzmeister, Felix; Kirchler, Michael; Razen, Michael (2019) : Cognitive skills and economic preferences in the fund industry, Working Papers in Economics and Statistics, No. 2019-16, University of Innsbruck, Research Platform Empirical and Experimental Economics (eeecon), Innsbruck

This Version is available at:

<https://hdl.handle.net/10419/207082>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Cognitive Skills and Economic Preferences in the Fund Industry

**Adam Farago, Martin Holmen, Felix Holzmeister, Michael
Kirchler, Michael Razen**

Working Papers in Economics and Statistics

2019-16



University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
research platform "Empirical and Experimental Economics"
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
<https://www.uibk.ac.at/eeecon/wopec/>

For a list of recent papers see the backpages of this paper.

Cognitive Skills and Economic Preferences in the Fund Industry

Adam Farago[†] Martin Holmén[†] Felix Holzmeister[‡]
Michael Kirchler^{†,‡,*} Michael Razen[‡]

[†] University of Gothenburg, Department of Economics, Centre for Finance

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

* Corresponding author: michael.kirchler@uibk.ac.at

Abstract

By running a battery of incentivized and non-incentivized experiments with fund managers from four countries in the European Union, we investigate the impact of fund managers' cognitive skills and economic preferences on the dynamics of the mutual funds they manage. First, we find that fund managers' risk tolerance positively correlates with fund risk when accounting for fund benchmark, fund category, and other controls. Second, we show that fund managers' ambiguity tolerance positively correlates with the funds' tracking error from the benchmark. Finally, we report that cognitive skills do not explain fund performance in terms of excess returns. However, we do find that fund managers with high cognitive reflection abilities generate these returns at lower risk.

JEL: C91, D91, G11, G41, J24

Keywords: Cognitive skills, economic preferences, fund managers, fund performance, experimental finance.

We thank Peter Bossaerts, Matthias Heinz, Jürgen Huber, Matthias Stefan, Matthias Sutter, seminar participants of the C-SEB seminar at the University of Cologne, the Max Planck Institute for Research on Collective Goods in Bonn, Goethe University Frankfurt, as well as conference participants of the NOEG 2019 (Austrian Economic Association) in Graz and Experimental Finance 2019 in Copenhagen for valuable comments. We particularly thank all participating fund managers for the excellent collaboration and their enthusiasm. Financial support from the Austrian Science Fund FWF (START-grant Y617-G11 and SFB F63) and the Swedish Research Council (grant 2015-01713) is gratefully acknowledged. Depersonalized experimental data and the software used in the online experiment are available via <https://osf.io/4hgct/>. A demo version of the experiment including all instructions is available via <https://fea-2018-en.herokuapp.com>. This study was ethically approved by the Ethical Review Board in Gothenburg.

1. Introduction

There is a growing survey-based, experimental, and empirical literature showing that economic preferences influence investment decisions and portfolio returns among private investors. For instance, risk aversion is negatively related to portfolio underdiversification (Dorn and Huberman, 2005) and ambiguity aversion is positively correlated with portfolio risk (Bianchi and Tallon, 2018). Not only economic preferences, but also cognitive skills have been shown to drive private investors' behavior and decisions. For example, skills in financial literacy are positively correlated with (risk-adjusted) portfolio returns (Bianchi, 2018; Von Gaudecker, 2015) and various forms of cognitive skills predict traders' performance in laboratory asset markets (Corgnet et al., 2018).

Up to now, it is an open question whether similar behavioral patterns hold for institutional investors like fund managers as well, or whether their education, decade-long experience in the industry, incentives, and the institutional framework reduce or even eliminate the impact of economic preferences, personality traits, and cognitive skills on their professional behavior. For instance, attitudes towards risks and losses might not matter much, as fund managers are benchmarked on certain indices in the fund prospectus, leaving little room for deviating strongly from the benchmark. Moreover, fund managers' remuneration is typically a convex function of past performance relative to other funds, offering tournament incentives that might override individual-level preferences for risk and loss. In addition, cognitive skills and fund managers' willingness to compete might be determinants of (risk-adjusted) performance and fund risk, yet the high level of efficiency of financial markets might render their relative impact negligible. Given their central role in the finance industry, it is important to know more about how fund managers' preferences and cognitive skills shape their decisions.¹ Evidence from lab-in-the-field experiments with finance professionals offers first hints on the direction of potential effects. These results reveal that professionals' behavior is not necessarily closer to rational benchmarks compared to the behavior of standard laboratory subjects like students (e.g., Haigh and List, 2005; Kirchler et al., 2018b).

Monetary and non-monetary incentives (Kirchler et al., 2018b), the fund managers' institutional framework (Chevalier and Ellison, 1999), and the competitive pressure of the market make it unclear whether individual-level economic preferences and cognitive skills also affect their professional decisions. To overcome this problem, we investigate the link between fund managers' preferences and skills, measured in online experiments, and the real-world consequences of their professional behavior in terms of fund performance and risk.

In particular, we address the following research questions:

RQ1: Do cognitive skills like fluid intelligence, cognitive reflection, and Theory of Mind influence fund managers' performance and fund risk?

RQ2: Do economic preferences towards risk, losses, and ambiguous outcomes impact fund managers' performance and fund risk?

¹ For instance, the value of total net assets of international open-end funds was \$49.3 trillion in 2017. Moreover, 114,000 open-end funds were registered, amounting to \$2.7 trillion net sales in 2017. These numbers impressively underline the central role of fund managers for the industry and the economy in general (see http://www.icifactbook.org/data/18_fb_data).

In a first step, 84 fund managers from four large and mid-sized countries in the European Union (EU), managing 267 mutual funds with average assets under management (*AUM*) of €342 million, took part in our online experiment. Then, we compiled daily and monthly time series data of the managed funds and their associated benchmarks, including, among others, returns, *AUM*, and total expense ratios (*ER*). Finally, we matched the experimental data of each fund manager with their monthly funds' data and control for each fund's benchmark, the funds' category, time fixed effects, and other controls like tenure.

Specifically, we followed the literature and ran three standard, non-incentivized tests to measure cognitive skills (Bruguier et al., 2010; Corgnet et al., 2018). To obtain scores in fluid intelligence, we administered a test similar to Corgnet et al. (2018). The test consisted of 18 of the Raven's advanced progressive matrices (Raven, 2000) in which fund managers had to solve diagrammatic puzzles. For cognitive reflection skills, we used the extended version of the cognitive reflection test (CRT, see Frederick, 2005) from Primi et al. (2015) and Toplak et al. (2014). The questions are constructed such that they have an intuitive, but on reflection incorrect, response put forward by System 1; whereas the correct response requires the effortful activation of System 2 (Dual Process Theory, see Kahneman, 2011). For measuring Theory of Mind skills (TOM), which are important in detecting the informational content of investing by inferring others' intentions (Bruguier et al., 2010), we ran the "Reading-the-Mind-in-the-Eyes"-test proposed by Baron-Cohen et al. (2001).

Moreover, we ran four incentivized, high-stakes experiments eliciting risk preferences, attitudes towards losses, intertemporal preferences, and ambiguity tolerance. Following Falk et al. (2018), we measured risk by applying a staircase procedure where fund managers had to make multiple decisions between a risky lottery and a (varying) safe payment. To get a measure of loss tolerance, we ran a similar experiment with mixed lotteries, varying the potential losses while holding the alternative safe payment fixed. To test ambiguity tolerance, we applied a modified version of the ambiguity experiment of Dimmock et al. (2016). Here, subjects could choose between ambiguous and risky lotteries with varying probabilities of the latter.

Finally, alongside general demographic questions on age, experience in the industry, gender, and education, we added the five-item competition sub-scale of the Work and Family Orientation (wofo) questionnaire introduced by Helmreich and Spence (1978), measuring preferences to compete in general.

First, we find a strong positive relationship between fund managers' risk tolerance and fund volatility. This indicates that fund managers with lower (higher) levels of risk tolerance compose funds with lower (higher) fund volatility. Importantly, this finding holds while controlling for fund managers' self-selection into fund categories (i.e., fixed income, international equity) and for additional variables like fund benchmark, industry experience, and fund size. In economic terms, we observe that fund managers identified as risk-seeking in the experiment increase relative fund risk by up to 20.3 percentage points of the benchmark's volatility compared to risk-averse fund managers.

Second, we observe that fund managers with lower levels of ambiguity tolerance manage their funds with lower tracking errors compared to managers with higher levels of ambiguity tolerance. This finding indicates that tolerance towards ambiguous outcomes explains fund managers' propensity towards and tolerance of deviations from the benchmark.

Finally, we report that cognitive skills do not explain fund performance in terms of excess returns. However, we do find that fund managers with high cognitive reflection abilities generate returns at lower risk. In economic terms, we observe that relative fund risk of those fund managers with the highest CRT score is up to 12.4 percentage points lower than the one of fund managers with average and low CRT scores.

Our paper contributes to several strands of literature. We add to the strand of literature analyzing the interplay of economic preferences and portfolio dynamics among private investors (Bianchi and Tallon, 2018; Dimmock et al., 2016; Dorn and Huberman, 2005). To the best of our knowledge, there is only one study by Bodnaruk and Simonov (2016) focusing on professional investors (fund managers) applying an approach closely related to ours. The authors run a short and non-incentivized survey to elicit fund managers' risk and loss aversion. They show that fund managers with high levels of self-reported loss aversion in the survey construct mutual funds with lower downside risk, show lower fund performance, and have shorter careers in the finance industry. We contribute with a comprehensive approach focusing on cognitive skills and economic preferences we elicit in incentivized and non-incentivized economic experiments. We consider it remarkable that risk preferences and cognitive skills explain fund risk and risk-adjusted performance, given the incentives (Kirchler et al., 2018b), the institutional framework in the industry (Chevalier and Ellison, 1999), and the competitive pressure of the market, that could reduce or eliminate the role of professionals' preferences.

Second, we account for the literature investigating the impact of cognitive skills on investment behavior and performance. Here, a substantial body of work analyzes private investors' behavior outlining that high-IQ investors show higher levels of stock market participation (Christelis et al., 2010), earn higher Sharpe ratios (Grinblatt et al., 2011), are less prone to the disposition effect, and exhibit superior market timing and stock-picking skills, resulting in outperformance compared to low-IQ investors (Grinblatt et al., 2012). Experimental finance literature adds to these findings by analyzing the role of various cognitive skills for investment decisions: for instance, fluid intelligence measures the capacity to reason and solve novel problems and is necessary for logical problem solving (Mackintosh, 2011). Cognitive reflection adds to fluid intelligence because it measures the ability to engage in effortful reasoning (e.g., Oechssler et al., 2009; Toplak et al., 2011, 2014). Theory of Mind (TOM) defines one's capacity to infer others' intentions, which is considered important in detecting the informational content of trading by inferring others' strategies from order books and prices (Bruguier et al., 2010). Experimental evidence suggests that high cognitive reflection scores predict subjects' earnings in laboratory asset markets with student subjects (Corgnet et al., 2015b; Noussair et al., 2016), TOM correlates with subjects' skills in predicting price changes (Bruguier et al., 2010),² and all three concepts are joint predictors of trader performance (Corgnet et al., 2018). Moreover, Corgnet et al. (2015a) and Bosch-Rosa et al. (2018) find

² However, DeMartino et al. (2013) show that TOM skills can also be detrimental when trading on financial markets. In a study using fMRI techniques, the authors report a mechanism by which social signals affect value computations in ventromedial prefrontal cortex, thereby increasing subject's proneness to ride financial bubbles.

a causal relationship between traders' cognitive sophistication and price efficiency.³ We contribute by showing that cognitive skills drive (risk-adjusted) performance of fund managers and hence provide some external validity to laboratory results and findings among private investors.

Third, we contribute to the small but growing body of literature dealing with the behavior of financial professionals and to studies investigating personality traits and biases of business professionals in general (e.g., Adams et al., 2018; Graham et al., 2013; Kaplan et al., 2012; Malmendier and Tate, 2005). Regarding financial professionals, one major result is that their behavior can substantially differ from standard (student) subjects and representative general population samples. For instance, professionals exhibit a higher degree of myopic loss aversion (Haigh and List, 2005), are less prone to anchoring than students (Kaustia et al., 2008), can better discern the quality of public signals in information cascades (Alevy et al., 2007), and react more strongly to rank incentives (Kirchler et al., 2018b). However, professionals also show herding behavior similar to student subjects' (Cipriani and Guarino, 2009) and act in line with prospect theory (Abdellaoui et al., 2013). The key contribution of our study is that the battery of experimental tasks allows us to study the link between fund managers' economic preferences/cognitive skills and their professional decisions directly without resorting to indirect proxies or cohort effects.

Finally, we also add to the literature that examines whether mutual fund managers are "skilled". Carhart (1997), for example, concludes that there is no evidence of mutual fund managers being skilled based on the investors' net excess return (alpha). Similarly, Fama and French (2010) use alpha measures to obtain a cross-sectional distribution of managerial skill. They conclude that there is only superior performance—and hence indications of skill—in the extreme right tail of the distribution. On the other hand, there are studies suggesting that fund managers can indeed exhibit a persistent manifestation of skill. Berk and van Binsbergen (2015) use the value that a mutual fund extracts from the capital markets as a skill measure and document that some managers are skilled, and that skill may persist for up to ten years.⁴ This strand of literature infers manager skill and talent indirectly from various measures of fund performance and behavior. Our approach is fundamentally different and thus we contribute by estimating mutual fund managers' cognitive and non-cognitive skills directly with experimental tests and relate it to fund performance.

The remainder of the paper is organized as follows: In Section 2 we introduce the experimental design, the collection of the experimental and the empirical data, and our econometric approach. In Section 3 we present results and in Section 4 we discuss and conclude.

³ Another strand of literature explores the relationship between fund managers' observable characteristics such as age, experience, and education, respectively, and fund performance. For example, Golec (1996) documents a positive (negative) relation between fund performance and manager tenure and education (manager age). Chevalier and Ellison (1999) document a positive (negative) relation between fund performance and manager age and college SAT score (manager tenure). Both studies investigate the U.S. market and they also document a strong positive relation between a MBA degree and fund performance. Fang and Wang (2014) investigate the Chinese market and also find that fund managers with a MBA degree perform better. They also report that Chartered Financial Analysts (CFA) perform better. Other fund manager characteristics have weak or unclear relations to fund performance.

⁴ Several other studies also suggest that at least some managers are skilled (see, for instance, Cremers and Petajisto, 2009; Grinblatt and Titman, 1989, 1993; Wermers, 2000).

2. Experimental Design

We exploit two sources of data: first, we collected data on cognitive skills and economic preferences by means of online experiments and, second, we matched the empirical fund time series with the experimental data of the fund managers who participated in the experiment. To ensure anonymity, fund managers' identities were replaced by randomly generated unique identifiers to match the experimental data with depersonalized, empirical fund data from various databases.

2.1. Experimental Fund Manager Data

We contacted approximately 900 fund managers via hard-copy letters and/or e-mails in which the study was outlined and which included personalized login credentials for participation in the online experiment.⁵ Fund managers were informed about the anonymous matching of the experimental data with the corresponding fund data. With the decision to participate, fund managers acknowledged to accept the informed consent of the experiment.

The experimental tasks were divided into three parts: (i) cognitive skills, (ii) incentivized economic preferences, and (iii) personality traits. Importantly, fund managers did not receive immediate feedback after each task, but were told in advance that they can select whether they want to receive feedback and background information on the experimental tasks after data collection has been completed. This was done to provide additional incentives to the fund managers to participate and to provide full disclosure of the background of the experiment. Upon completion of the experiment, one of the incentivized tasks was randomly chosen to determine the subjects' payout. In addition to their earnings from the corresponding task, subjects received a fixed participation fee of €25. Details on the experimental procedure and the feedback can be found in Appendix A.

First, to measure fund managers' cognitive skills, we administered three different tasks. For cognitive reflection skills, we compiled a set of five questions taken from the extended cognitive reflection tests (CRT) proposed by Toplak et al. (2014) and Primi et al. (2015). The concept of cognitive reflection rests upon the dual-process theory framework (Kahneman, 2011). The questions of cognitive reflection tests are constructed in a way that they have an intuitive, but on reflection incorrect, response put forward by System 1; the correct response requires the effortful activation of System 2.⁶ To obtain a test score for fluid intelligence, we conducted a task similar to Corgnet et al. (2018), presenting 18 items of the Raven's Advanced Progressive Matrices (APM; Raven, 2000). In each of these items, subjects have to recognize the geometric pattern in an unfinished diagrammatic puzzle and identify the missing element. The main objective of this test is to measure subjects' ability to solve novel problems, which is why it is also used to measure IQ. One additional advantage is that it can discriminate well even among high-IQ subjects. To measure Theory of Mind skills (TOM), we used 18 pictures of the "Reading-the-Eyes-in-the-Mind"-test proposed by Baron-Cohen et al. (2001). In this test, subjects are shown photographs of

⁵ The total number of 900 invitations includes undelivered and returned mails, bounceback e-mails, outdated or invalid (e-mail) addresses, etc. Thus, the response rate of roughly 10% should be considered being a conservative lower bound.

⁶ For illustrative purposes, this is one of the questions: "Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?" (Toplak et al., 2014). The (incorrect) intuitive answer (30 students) can be "overruled" upon reflection (29 students), which requires effortful System 2 processes.

the eye region of different people and choose one of four feelings that best describe the mental state of the person whose eyes are shown. This test measures one's capacity to infer others' intentions, which, for instance, is important in detecting the information disseminated by the behavior of other market participants (Bruguier et al., 2010).⁷

Second, to measure economic preferences of the fund managers, we administered four incentivized experiments. Subjects were informed that, at the end of the experiment, one of these tasks would be randomly chosen and their decision in the respective task would determine their payout. Risk attitudes and inter-temporal preferences were elicited as in Falk et al. (2018).⁸ The task for loss tolerance was adapted from the procedure of Gächter et al. (2007), while ambiguity tolerance was measured following the design introduced by Dimmock et al. (2016). We increased consistency and comparability of the experiments by presenting all tasks in a staircase framework (see Figure S2 in Appendix A.2 for one example following Falk et al., 2018). In this setting, subjects face a set of path dependent decisions, offering two choices each. Along these decisions, one option stays the same, while the second option depends on the previous choice. Compared to single and multiple price list formats, this procedure offers the advantage to be concise without forfeiting precision in eliciting points of indifference.

In the risk preferences task, subjects first had to choose between a lottery paying €60 or €0 with equal probability and a safe payment of €32. Subjects who preferred the lottery in the first stage were presented a higher safe payment in the second stage, while subjects who preferred the safe payment were presented a lower safe payment in the second stage. After four stages, this design allows to pin down a narrow interval for the subjects' certainty equivalents and hence an estimate of their risk preferences. Clearly, those subjects with high certainty equivalents are considered to show high levels of risk-tolerance. The payout of the safe alternative varied from €4 to €60.⁹

The loss aversion task started with the question whether subjects preferred to participate in a lottery that pays €22 or €-12 with equal probability. The positive payoff of €22 stayed the same in all questions. Subjects who rejected the lottery were presented with a lower negative payoff in the second stage while subjects who accepted the lottery were presented with a higher potential loss in the second stage. Iterating this procedure reveals the maximum loss subjects were willing to accept in order to obtain the chance of winning €22. According to this logic, subjects with a high tolerable maximum loss are the ones with high levels of loss tolerance. The range of varying negative payouts in the lottery varied from €-22.50 to €-1.50.¹⁰

⁷ For the Raven's Advanced Progressive Matrices (APM) and the eye-gaze test (TOM), we used shortened versions. The original tasks comprise 36 questions each, out of which we took every second question, starting with the first one of the original task. This was done to keep the overall time needed to complete the survey as short as possible without losing explanatory power.

⁸ For the main analyses, we had to drop the variable on inter-temporal preferences (PATIENCE). The reason is that there was a runtime error issue with this task because of which we lost part of the observations on PATIENCE, leaving us with an unbalanced panel otherwise. However, analyses of the sub-sample reported in Table S8 in Appendix C show no impact of PATIENCE on any of our dependent fund variables.

⁹ In the dropped time preferences task, the first decision problem asked subjects whether they preferred a payment of €20 today or a payment of €31 in 6 months. Those who selected the payment today were presented a higher future payment in the second stage while those who went with the future payment were presented a lower future payment in the second stage. Iterating this procedure reveals the implicit time discounting rate of the subjects. Subjects who predominantly select the future payment are the ones with high levels of patience (or, in other words, lower time discounting rates).

¹⁰ Note that this parametrization together with the participation fee of €25 ensured that subjects could not incur losses in the experiment.

Table 1: Descriptive summary statistics of scores in the experimental tasks.

<i>Task (Variable)</i>	Mean	SD	Min	Max
CRT	3.95	1.09	1	5
APM	9.81	2.85	1	16
TOM	11.51	2.51	4	16
RISK TOLERANCE	27.14	6.76	10	50
LOSS TOLERANCE	18.05	5.55	2.25	23.25
AMBIGUITY TOLERANCE	39.83	13.92	4	78
COMPETITIVENESS	26.61	4.40	15	35

Note: The cognitive reflection test consisted of 5 questions (CRT). The task measuring fluid intelligence comprised of 18 questions (APM). The "Reading-the-Eyes-in-the-Mind"-test measuring theory of mind comprised of 18 questions (TOM). The score for risk tolerance reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability (RISK TOLERANCE). The score for loss tolerance reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22 (LOSS TOLERANCE). The score for ambiguity tolerance represents the matching probability (in %) that left subjects indifferent between a risky lottery with the respective probability of winning and an ambiguous lottery with an unknown probability of winning (AMBIGUITY TOLERANCE; both lotteries paid €60 in the case of winning and €0 else). Competitiveness was measured as the sum of all five questions of the wofo survey with Likert-scales from 1 to 7 each (COMPETITIVENESS).

In the ambiguity preferences task, the first decision problem asked subjects to choose between two lotteries. Each lottery offered the chance to win €60 or €0. While the probability of winning €60 was known to be 50% in one of the lotteries (risk), it was unknown in the other lottery (ambiguity). The ambiguous lottery remained unchanged throughout the task. Subjects who chose the risky lottery were presented with a new risky lottery offering a lower known probability of winning in the second stage, while subjects who chose the ambiguous lottery were presented a new risky lottery offering a higher known probability of winning in the second stage. Iterating this procedure reveals the matching probability at which subjects are indifferent between the ambiguous and the risky lottery. Thus, subjects who predominantly select the ambiguous lottery are the ones with high levels of ambiguity tolerance. The probabilities for winning €60 in the risky lottery ranged from 7% to 93%.

Third, we ran a test on measuring fund managers' attitudes towards competition. We used the 5-item Work and Family Orientation (wofo) questionnaire proposed by Helmreich and Spence (1978), which is a widely used psychometric measure of individuals' self-assessed competitiveness and which was previously used in experiments with financial professionals (Kirchler et al., 2018a).¹¹

Questions on demographics concluded the experiment. In total, 94 fund managers completed the experiment. We were able to retrieve fund and benchmark data for 84 of these 94 (see section 2.2 for details). We lose one fund manager, because we were not able to match at least one fund to him/her, and we lose 9 out of the remaining 93 fund managers, because we are not able to obtain unambiguous benchmark data for any of the funds they manage. Consequently, our final results are based on a sample of 84

¹¹ Subjects answered the following five questions: "I enjoy working in situations involving competition with others"; "It is important to me to perform better than others on a task"; "I feel that winning is important in both work and games"; "It annoys me when other people perform better than I do"; "I try harder when I'm in competition with other people". The answers were provided on a Likert scale from 1 (I do not agree at all) to 5 (I fully agree). The sum over all five questions finally enters our data analyses.

fund managers. The average age of the 84 fund managers in our final sample was 44 years, with an average tenure in the finance industry of 18 years. 95% of the fund managers were male. The experiment was programmed in *oTree* (Chen et al., 2016), utilizing the ready-made applications introduced by Holzmeister (2017). Completing the online experiment took fund managers on average 32 minutes (*SD* of 9 minutes). Payout to the subjects was electively administered via a third party specialized on micro-payments or via bank transfer.

Table 1 provides a first descriptive overview of the experimental results. For the econometric analyses, we normalized the cognitive skill and competitiveness measures by subtracting means and dividing by standard deviations. This does not affect *t*-statistics, but it makes the economic interpretation of the corresponding coefficients more meaningful. For the economic preference tasks, we conduct our analysis using the original metrics of the elicitation procedures.¹²

2.2. Empirical Fund Data

We use *Morningstar Direct* to match the participants to empirical data on the funds they managed between January 2008 and December 2017. As indicated, using the procedure detailed in Appendix B.1, we are able to match at least one fund to 93 out of the 94 participants. The resulting sample contains 369 funds altogether for a total of 20,734 unique fund-month observations.

After identifying the funds, we collect time series data on their net returns both on the daily (denoted by R_{id}^{net}) and monthly (denoted by R_{it}^{net}) frequencies. For each fund, we also obtain the monthly total expense ratios, ER_{it} , and the series of end-of-month assets under management, AUM_{it} . The main source of the data is the *Morningstar Direct* database, but we augment the *ER* and *AUM* series using the *Lipper* database and imputations. Details of the data collection and imputations are described in Appendix B.2.

Using the monthly data, we calculate the following variables related to fund performance:

$$R_{it}^{gross} = R_{it}^{net} + ER_{it} \quad (1)$$

$$R_{it}^{abn} = R_{it}^{gross} - R_{it}^B \quad (2)$$

$$V_{it} = AUM_{it-1} \cdot R_{it}^{abn} \quad (3)$$

The gross return, R_{it}^{gross} , represents the total return on the fund before expenses. R_{it}^B denotes the monthly return on the fund's benchmark. Hence the abnormal return, R_{it}^{abn} , is the total monthly return of the fund over its benchmark. We also calculate the value added, V_{it} (Berk and van Binsbergen, 2015), i.e., the product of assets under management and abnormal returns. Berk and van Binsbergen (2015) argue that the skill of a mutual fund manager equals to the value her fund extracts from capital markets, which can be measured by V_{it} .

¹² The motivation for normalizing cognitive skill and competitiveness measures is the absence of interpretable economic units. Moreover, accounting for potential differences in scaling among the wofo questions, we also normalized each question separately before computing aggregated competitiveness scores.

To measure the amount of risk (relative to the benchmark) taken by the fund managers, we calculate the following two measures using daily data:

$$RV_{it} = \frac{Std\left(R_{id[t]}^{net}\right)}{Std\left(R_{id[t]}^B\right)} \quad (4)$$

$$TE_{it} = Std\left(R_{id[t]}^{net} - R_{id[t]}^B\right), \quad (5)$$

where $R_{id[t]}$ denotes daily returns within month t .¹³ The relative volatility, RV_{it} , measures the overall riskiness of the fund relative to the riskiness of the benchmark. A value of 1 indicates that, during month t , the fund’s return volatility was the same as that of the benchmark, while values above (below) one indicate that the return volatility of the fund was higher (lower) than that of the benchmark. The tracking error, TE_{it} , measures risk in the fund’s return that is due to active management decisions made by the portfolio manager. The two variables measure risk-taking from two different points of view: RV_{it} compares the overall level of risk taken by the fund and its benchmark, while TE_{it} indicates how closely the fund mimics its benchmark index.

Almost all of the above variables related to performance and risk-taking require return data on the fund’s benchmark, R_{it}^B . We use the prospectus benchmark, which is self-declared by the fund company itself, because this benchmark is likely to have the highest influence on the manager’s decisions. By comparing each fund to its prospectus benchmark, we follow the approach of, for instance, Fama and French (2010) and Ibert et al. (2018).¹⁴ However, our mutual fund databases do not report a prospectus benchmark for all funds in our sample. We are able to identify and obtain return data on the prospectus benchmark for 182 funds (49% of the 369 funds in our base sample). Since we would like to avoid losing half of our sample, we assign benchmarks to remaining funds, if possible, via two further steps. First, similar to Ibert et al., 2018, we use the following benchmark assignment rule for equity funds: for each equity fund category defined by the *Morningstar* variable “Category”, we find the most common benchmark among all open-ended mutual funds that have one of our four countries registered as “Domicile”. This most common benchmark is assigned to all the funds in the given category. We are able to obtain benchmark return data for 39 additional funds this way. Second, we use *Lipper* to identify further benchmarks. *Lipper* independently assigns the “Lipper Technical Indicator Benchmark” to most of the funds in the database according to its assessment of the fund’s investment strategy. This technical indicator benchmark is used for the remaining funds, if available. Benchmark return data is obtained for 46 additional funds via this step. Further details of the benchmark assignment process are described in Appendix B.3. Altogether, there are 267 funds with benchmark data and these funds provide the sample for our estimation results. Note that we lose 9 managers out of the 93 to whom we were able to match at least one fund, because we are not able to obtain unambiguous benchmark data for any of the

¹³ We require at least 18 daily observations to calculate the monthly standard deviations. Note also that using $R_{id[t]}^{gross}$ instead of $R_{id[t]}^{net}$ would lead to exactly the same results since the expense ratio, ER , is constant within the month.

¹⁴ Sensoy (2009) argues that due to strategic benchmark selection, the fund’s self-declared benchmark might not match the fund’s actual investment style. Therefore, we also use an alternative benchmark assignment rule in the robustness section, where we only use benchmarks assigned by Lipper based on its assessment of the given fund’s investment strategy. Another alternative would be to use a benchmark based on a factor model like the three-factor model of Fama and French (1993) or the four-factor model of Carhart (1997). However, given the large heterogeneity of the funds in our sample with respect to country of domicile, allocation across asset classes, and geographical focus, it is unclear what factors should be used and the risk of using a misspecified benchmark is high. Therefore, we do not pursue this approach.

funds they manage. Consequently, our results are based on a sample of 84 fund managers.

Fund-level summary statistics are provided in Table 2. Panel A corresponds to the final sample of 267 funds with benchmark return data, while Panel B presents the base sample of 369 funds. Comparison of the two panels reveals minor differences that are due to the availability of the benchmark return data. The following discussion focuses on the final sample (i.e., on Panel A), while the minor differences compared to the base sample are highlighted at the end of the section.

The average fund (out of the 267) appears in the sample through 53 months, which is roughly half of the full sample period of 10 years. The average fund has AUM of €342 million, but the cross-sectional fund size distribution is wide and left skewed: the median AUM is €78 million, ten percent of the funds are smaller than €6.8 million, while ten percent of the funds manage more than €958 million. The average annual expense ratio is 1.30%, but investors can pay as little as 0.41% (10th percentile) and as much as 2.04% (90th percentile). The average annual gross return is 6.58% (the median is 7.12%), and there is a large variation across funds with the interdecile range taking values from -0.57% to 15.12%. The funds earn a slightly higher gross return than their benchmark on average: the average (median) annual abnormal return is 1.04% (0.58%). However, the abnormal return after expenses ($R_i^{abn} - ER_i$) is negative both for the average and for the median fund (not reported in the table). The average annual value added is €6.4 million, which is much higher than the median of €0.7 million, indicating that the distribution of V_i is considerably right skewed. Both the average and median RV_i are close to one, indicating that the return volatility of a typical fund is close to the volatility of its benchmark. However, RV_i can be as low as 0.81 (10th percentile) indicating that the fund's volatility is 19% lower than that of the benchmark, or can be as high as 1.23 (90th percentile), indicating that the fund is 23% more volatile than its benchmark. Less than half of the funds (42%) are single-managed. The funds' distribution across major investment categories is 66% equity funds, 20% fixed income funds, 10% allocation funds, and 4% others.

The comparison of the two panels in Table 2 shows that the tenure, expense ratio, and return statistics are very similar in the base sample and the final sample. There are two differences between the two samples: (i) larger funds are over-represented in the final sample, as the average and median AUM is slightly higher in Panel A, and (ii) allocation funds are under-, while equity funds are over-represented in the final sample, relative to the base sample. This is due to the fact that smaller and allocation funds less frequently report a benchmark and that it is more difficult to come up with a suitable one.

2.3. Econometric Model

To answer our research questions on the impact of cognitive skills and economic preferences, we set up the following regression model. As dependent variables, indicating fund performance, we use monthly abnormal returns, R_{it}^{abn} , and monthly value added, V_{it} for each fund. As proxies for fund risk, we use relative volatility, RV_{it} , and tracking error, TE_{it} , serving as dependent variables as well.

As independent variables, we include all experimentally elicited variables measuring cognitive skills and we add fixed-effects for fund category and time. Moreover, we add a dichotomous variable indicating whether a fund was team-managed or not ($TEAM$) in the given month, $EXPERIENCE$ measuring years in the industry, and the log of assets under management at the end of the previous fund, $\log(AUM)_{t-1}$. In

Table 2: Summary statistics at the fund level.

Panel A: Funds with benchmark return data								
	N	$Mean$	SD	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}
in sample (months)	267	53.34	36.51	12.00	24.00	44.00	84.00	117.00
AUM_i (mio. EUR)	267	342.26	710.20	6.76	25.95	78.31	289.82	957.81
ER_i (annual %)	267	1.30	0.90	0.41	0.73	1.33	1.69	2.04
R_i^{gross} (annual %)	267	6.58	10.39	-0.57	2.97	7.12	12.13	15.12
R_i^{net} (annual %)	267	5.28	10.40	-1.18	1.64	5.64	10.73	14.05
R_i^{abn} (annual %)	267	1.04	4.45	-2.99	-0.80	0.58	2.65	5.38
V_i (annual mio. EUR)	267	6.39	31.79	-3.13	-0.36	0.25	2.22	11.79
RV_i	263	1.03	0.26	0.81	0.95	1.01	1.09	1.23
TE_i	265	0.71	0.55	0.14	0.30	0.55	1.06	1.38
Team-managed	267	0.58						
Fixed income	267	0.20						
Equity	267	0.66						
Allocation	267	0.10						
Rest	267	0.04						
Panel B: All funds								
	N	$Mean$	SD	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}
in sample (months)	369	56.19	37.59	10.00	26.00	49.00	88.00	120.00
AUM_i (mio. EUR)	364	281.11	631.25	6.76	18.51	58.26	220.88	774.93
ER_i (annual %)	362	1.27	0.82	0.42	0.74	1.28	1.65	2.02
R_i^{gross} (annual %)	362	6.08	9.63	0.38	2.94	5.83	10.92	14.58
R_i^{net} (annual %)	369	4.76	9.58	-1.08	1.67	4.54	9.44	13.02
Team-managed	369	0.57						
Fixed income	369	0.17						
Equity	369	0.49						
Allocation	369	0.29						
Rest	369	0.05						

Note: The gross return, R_i^{gross} , represents the total return on the fund before expenses. R_i^B denotes the annual return on the fund's benchmark. The abnormal return, R_i^{abn} , is the total annual return of the fund over its benchmark. The value added Berk and van Binsbergen (2015), V_i , is the product of assets under management, AUM_i , and the abnormal return, R_i^{abn} . The relative volatility, RV_i , measures the overall riskiness of the fund relative to the riskiness of the benchmark, based on daily data ($R_{i,d[t]}$ denotes the daily net returns within month t). The tracking error, TE_i , measures risk in the fund's return that is due to active management decisions made by the portfolio manager.

the regressions regarding risk-taking, we also control for past fund performance. As it will be evident from the results, it is important to allow for a non-linear effect of past performance. Therefore, we use two variables that are non-linear transformations of the past fund return: $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to the gross return in the previous month if it was negative. Similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to the gross return in the previous month if it was positive. Including these two variables allows for the possibility to estimate a V-shaped effect of past return.

We use time fixed effects (i.e., monthly) and investment category fixed effects (i.e., the four categories shown in the last four rows of Table 2) in all our regressions. Standard errors are clustered at the manager level. Importantly, we had to drop our measure of fluid intelligence (APM) in most specifications due to multicollinearity with CRT.¹⁵ In general, APM yields qualitatively similar results in most specifications (when dropping CRT instead) and we mention results on APM where necessary. Finally, we would like to emphasize that we lower the p -value thresholds for statistical significance to 5.0% and 0.5% in all econometric specifications to reduce the likelihood of false positives and hence to improve the robustness of scientific findings.

In general, (i) we test for linear effects in all independent variables jointly in a first step. In additional robustness checks, (ii) we add the sub-sample of single-managed funds and (iii) we split up significant, independent variables found in step (i) to measure potential non-linear effects in the explanatory variables.

The matching of experimentally elicited cognitive skills and economic preferences with fund time series of the last ten years warrants some discussion. First, one alternative would be to run the experimental tasks in a first step and use it as predictors of fund performance and risk in the upcoming years. We leave this issue for future research, as we will analyze whether cognitive skills and economic preferences actually predicted (risk-adjusted) abnormal returns, fund risk, and tracking error in five to ten years. However, Bodnaruk and Simonov (2016) used the survey questions also for their matched sample of both approaches (i.e., the empirical time series were selected from the time span before and after the survey was run) and they do report similar results. Second, cognitive skills appear to be relatively stable over time, making the issue of experiments lagging or leading the empirical data less important. For instance, Böhm et al. (2018) use cognitive scores from the military enrollment tests at the age of 18 as a predictor for successful careers in finance. Moreover, a substantial body of literature uses the same approach of running experiments or surveys on economic preferences and relate its findings to portfolio choice of the preceding years (e.g., Bianchi, 2018; Bianchi and Tallon, 2018; Dimmock et al., 2016; Riedl and Smeets, 2017). Hence, these studies implicitly assume relative stable preferences which is also partly backed up by literature (Meier and Sprenger, 2015).

¹⁵ See Table S1 in Appendix A.2 for further details on the correlation between the experimental measures.

3. Results

Result 1: *Neither cognitive skills nor economic preferences and attitudes towards competition do systematically contribute to abnormal returns or value added.*

Support: As outlined in the first (R^{abn}) and third (V) column of Table 3, we observe that all coefficients of all experimental measures are insignificant. This suggests that neither fund managers' attitudes towards risk, losses, and ambiguity, nor their cognitive skills or their attitudes towards competition have an impact on the managed funds' abnormal returns or value added. Positive coefficients for the TEAM dummies ($p < 0.05$), however, suggest that, on average, team managed funds outperform single-managed funds by 0.1 percentage points per month and add €267,000 more to funds' value than single-managed funds on a monthly basis.

When turning to the subsample of single-managed funds, we find no effect of cognitive skills (CRT and TOM) on outperformance either, but we do find an impact of risk attitudes and ambiguity tolerance on abnormal returns. In particular, we report a negative relationship between risk tolerance and abnormal returns, indicating that more risk-tolerant fund managers underperform fund managers with lower levels of risk tolerance. In addition, we find a positive effect of ambiguity tolerance on fund performance. Taken together, this implies that less risk tolerant and more ambiguity-tolerant fund managers on average perform better than their counterparts in single-managed funds. These results, however, should not be overemphasized at this stage. We will discuss the latter finding together with the results on tracking errors outlined in Table 5.

Result 2: *Fund managers' risk tolerance is positively correlated with fund volatility. At the same time, fund managers with high cognitive reflection abilities take fewer risk.*

Support: As shown in columns 1 and 3 of Table 4, we observe that the coefficients of RISK TOLERANCE are significantly positive. This result also holds for fund managers of single-managed funds, even with a higher magnitude. This finding indicates that individual's risk preferences, measured in our lottery experiment, explain the riskiness of a managed fund, even beyond controls like the funds' benchmark and category, assets under management, time fixed-effects, years in industry, and the other experimental measures. With this procedure we also control for potential effects of self-selection of fund managers into certain fund categories (e.g., fixed income vs. equity funds) and find the effect of risk preferences on top of the controls. Moreover, when decomposing the effect of risk preferences on relative fund volatility in columns 2 and 4, we observe that it is robust for the classification along risk-seeking (certainty equivalent exceeds the expected value of the lottery), risk-neutrality, and risk-aversion (expected value exceeds the certainty equivalent). This means that, for instance, risk-seeking fund managers compose funds with significantly higher levels of fund risk compared to their risk-neutral or risk-averse counterparts. We also consider the economic significance to be important. As outlined in columns 2 and 4, risk-seeking fund managers increase relative fund risk by 15.6% and 20.3%, respectively, of the benchmark's volatility compared to those of risk-averse fund managers.

Besides gaining insights on drivers of fund performance and risk, we also interpret our study as a test for the external validity of experimentally elicited economic preferences. Our non-significant finding on

Table 3: Performance – abnormal returns and value added: The table shows the results of ordinary least squares regressions of funds’ abnormal returns and value added on cognitive skills and economic preferences/attitudes. The abnormal returns, R^{abn} , are the funds’ monthly gross returns over their benchmark. The value added, V , is the product of assets under management and abnormal returns. Standard errors are clustered at the manager level. Corresponding p -values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

<i>Dependent variable</i> <i>Sample of funds</i>	<i>Abnormal returns</i>		<i>Value added</i>	
	R^{abn} <i>All</i>	R^{abn} <i>Single</i>	V <i>All</i>	V <i>Single</i>
CRT	−0.015 (0.497)	−0.032 (0.172)	0.035 (0.570)	−0.040 (0.506)
TOM	−0.013 (0.555)	−0.008 (0.765)	−0.014 (0.761)	−0.040 (0.260)
COMPETITIVENESS	0.032 (0.256)	0.024 (0.560)	0.062 (0.369)	0.071 (0.147)
RISK TOLERANCE	−0.001 (0.813)	−0.011* (0.008)	0.007 (0.515)	−0.004 (0.389)
LOSS TOLERANCE	−0.002 (0.670)	−0.001 (0.741)	−0.004 (0.685)	−0.003 (0.625)
AMBIGUITY TOLERANCE	0.003 (0.063)	0.006** (0.003)	0.003 (0.283)	0.005 (0.153)
$\log(AUM)_{t-1}$	0.016 (0.167)	−0.007 (0.696)		
EXPERIENCE	0.001 (0.602)	0.000 (0.904)	0.001 (0.898)	0.005 (0.390)
TEAM	0.099* (0.010)		0.267* (0.019)	
Constant	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	14,625	7,179	14,625	7,179
Number of managers	84	53	84	53
Adjusted R^2	0.124	0.124	0.076	0.072

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the “Reading-the-Mind-in-the-Eyes”-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each. $\log(AUM)$ stands for the log of assets under management, EXPERIENCE indicates years in industry, and TEAM is a dichotomous indicator for team-managed funds.

loss aversion and fund risk, however, is in contrast to the result of Bodnaruk and Simonov (2016), showing that fund managers who exhibit high levels of loss aversion in a non-incentivized survey construct funds with lower downside risk and exhibit lower fund performance. We contribute with a comprehensive approach on economic preferences with incentivized economic experiments and do not find any impact of loss aversion on fund volatility, but we do find a strong impact of experimentally elicited risk preferences on fund risk. This also hints at a high external validity of experimentally elicited risk preferences.

In addition, our analysis provides evidence that CRT scores are negatively related to relative fund volatility. As outlined in the first and third column of Table 4, we report that the coefficients of CRT are negative (at the 5%-level) for team- and single-managed funds alike. Decomposing this effect in columns 2 and 4 reveals its non-linear nature, as it is actually caused by the fund managers with the highest score (see the binary variable $\text{CRT} = 5$ for those fund managers who answered all five questions correctly).¹⁶ Again, we argue that this effect is also economically relevant. In comparison to fund managers with CRT scores of 3 or less, relative fund risk for those fund managers with the highest possible CRT score is 9.5 and 12.4 (single-managed funds) percent lower (see columns 2 and 4).

Turning to the control variables, we report that fund managers' risk-taking is strongly driven by previous month's return realizations, and the effect is non-linear. We observe significantly negative coefficients for $\min(R_{t-1}^{gross}, 0)$ (which can take on values ≤ 0) and significantly positive coefficients for $\max(R_{t-1}^{gross}, 0)$ (which can take on values ≥ 0). This indicates that both larger negative and larger positive returns in the previous month lead to higher risk-taking relative to the benchmark. We hence report a V-shaped reaction of risk-taking on past performance. Let us highlight two further observations regarding past performance as a control. First, it is important to model the non-linearity of the effect; if past performance is included simply as a linear variable instead of the two non-linear variables above, its coefficient is insignificant. Second, if we do not control for past performance (i.e., leave out $\min(R_{t-1}^{gross}, 0)$ and $\max(R_{t-1}^{gross}, 0)$ from the regressions) or control for past performance linearly, the coefficients on all remaining explanatory variables are effectively unchanged, hinting at robust patterns of the behavioral variables. We find no impact of all the other control variables on the relative riskiness of the fund to its benchmark.

Finally, taking *Result 1* and *Result 2* on cognitive skills together, we summarize that fund managers with higher levels of cognitive reflection skills (CRT) manage to obtain the same returns at considerably lower risk than fund managers with lower CRT scores. Referring to our interpretation of the study as a test for external validity of laboratory experiments, we contribute with our results on the interplay of cognitive skills and (risk-adjusted) investment performance. As outlined in section 1, experimental evidence suggests that high cognitive reflection scores predict subjects' earnings in asset markets with student subjects (Corgnet et al., 2015b; Noussair et al., 2016), TOM correlates with subjects' skills in predicting price changes (Bruguier et al., 2010), and all three concepts (APM, CRT, and TOM) are joint predictors of trader performance (Corgnet et al., 2018). Moreover, Chevalier and Ellison (1999) document a positive relationship between risk-adjusted fund performance and college SAT scores by drawing on

¹⁶ We run the same regressions as reported in columns 1 and 3 of Table 4 with APM instead of CRT. We find similar, but somewhat weaker effects. In particular, we observe a significant and non-linear effect of APM on relative volatility, as the top fund managers managing to solve 13 or more out of the 18 questions compose their funds with significantly lower relative fund volatility compared to their peers.

Table 4: Risk – relative volatility: The table shows the results of ordinary least squares regressions of funds’ relative volatility on cognitive skills and economic preferences/attitudes. The relative volatility, RV , measures the overall riskiness of the fund relative to the riskiness of the benchmark. Standard errors are clustered at the manager level. Corresponding p -values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

<i>Dependent variable</i> <i>Sample of funds</i>	<i>Relative volatility</i>			
	<i>RV</i> <i>All</i>	<i>RV</i> <i>All</i>	<i>RV</i> <i>Single</i>	<i>RV</i> <i>Single</i>
CRT	-0.028* (0.024)		-0.036* (0.039)	
CRT = 4		0.005 (0.884)		0.050 (0.204)
CRT = 5		-0.095* (0.013)		-0.124** (0.005)
TOM	-0.010 (0.629)	-0.003 (0.899)	-0.050* (0.007)	-0.035 (0.074)
COMPETITIVENESS	0.016 (0.265)	0.017 (0.268)	0.011 (0.570)	-0.008 (0.716)
RISK TOLERANCE	0.006** (< 0.001)		0.010** (0.001)	
RISK NEUTRAL		0.058 (0.083)		0.090* (0.027)
RISK SEEKING		0.156** (0.001)		0.203* (0.036)
LOSS TOLERANCE	0.000 (0.925)	0.000 (0.983)	0.002 (0.622)	0.003 (0.409)
AMBIGUITY TOLERANCE	-0.002 (0.171)	-0.002 (0.162)	-0.002 (0.086)	-0.002 (0.267)
$\min(R_{t-1}^{gross}, 0)$	-0.011** (0.001)	-0.012** (< 0.001)	-0.010* (0.013)	-0.011** (0.004)
$\max(R_{t-1}^{gross}, 0)$	0.010** (< 0.001)	0.012** (< 0.001)	0.013** (< 0.001)	0.015** (< 0.001)
Constant & Controls	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	14,183	14,183	6,784	6,784
Number of managers	84	84	53	53
Adjusted R^2	0.126	0.138	0.230	0.259

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the “Reading-the-Mind-in-the-Eyes”-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the wofo, answered on scales ranging from 1 to 7 each. $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month’s gross return if it was negative; similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month’s gross return if it was positive. Controls refer to lagged log of assets under management ($\log(AUM)_{t-1}$), years in industry (EXPERIENCE), and a dichotomous indicator for team-managed funds (TEAM). All controls are insignificant.

empirical data. We add mixed evidence to these findings. On the one hand we do not find a clear impact of the various cognitive skills on individual performance, as shown in the above cited experiments. In particular, we do not find any impact of Theory of Mind skills on performance. On the other hand we do find that cognitive reflection abilities do explain performance when accounting for mutual funds' risks. Hence, we contribute by showing that high CRT fund managers generate the same fund return, but at significantly lower fund risk.

Result 3: *Fund managers' preferences for ambiguous outcomes are positively correlated with funds' tracking errors.*

Support: As shown in the first column of Table 5, we observe a pattern of ambiguity tolerance being positively related to the managed funds' tracking error. While this pattern is suggestive for the sample of all funds, it turns out to be statistically significant for single-managed funds (column 3 in Table 5). This finding indicates that when fund managers are more tolerant towards ambiguous outcomes they manage funds that deviate stronger from their benchmark compared to fund managers with lower levels of ambiguity tolerance. In addition, when decomposing this finding into ambiguity neutral and ambiguity seeking fund managers, we show the robustness of the findings: ambiguity-seeking fund managers, on average, manage their funds with considerably higher tracking errors compared to their ambiguity-averse peers. In economic terms, we find that ambiguity seeking fund managers exhibit a monthly tracking error that is up to 0.354 percentage points larger than the one of ambiguity-averse fund managers (see column 4 in Table 5). Given that the average monthly tracking error in our sample is 0.71%, the effect of ambiguity tolerance seems quite remarkable.

Similar to patterns identified for the funds' relative volatility (Table 4), we report that fund managers' tracking errors are dependent on previous months' return realizations. We observe significantly negative coefficients for $\min(R_{t-1}^{gross}, 0)$, indicating that fund managers' tracking errors increase with increasing negative gross returns in the previous month. In addition, we find a similar, yet less pronounced effect on tracking errors of larger positive past gross returns. As for relative volatility, this hints at a V-shaped pattern of past return-based tracking errors with a stronger effect on the downside. Again, including a linear variable for past performance or removing both variables leaves the effects of the remaining explanatory variables unchanged.

Taking the results on relative volatility and tracking error together, our findings suggest that fund managers facing negative gross returns in the previous month subsequently increase the risk of their portfolios, which in turn leads to higher tracking errors, i.e. higher deviations from the benchmark. This brings us back to the initial finding on Table 3 concerning the positive relationship of ambiguity tolerance and performance of single-managed funds. If ambiguity tolerance works via the risk channel (relative volatility and tracking error), it could positively affect abnormal returns during phases of generally rising stock markets, as in our sample. We, however, conjecture that the impact of tracking errors on abnormal returns could also be different for other time periods and therefore we interpret this finding in Table 3 with great caution.

Moreover, we find a negative impact of fund size, $\log(AUM)_{t-1}$, on the tracking error, indicating that larger funds are managed with lower tracking errors. Finally, we report no clear pattern regarding the impact of the other control variables on the relative riskiness of the fund to its benchmark.

Table 5: Tracking error: The table shows the results of ordinary least squares regressions of funds' tracking errors from the benchmark on cognitive skills and economic preferences/attitudes. The tracking error, TE , measures risk in the fund's return that is due to active management decisions, i.e., fund's net return minus the benchmark return. Standard errors are clustered at the manager level. Corresponding p -values are shown in parentheses. ** $p < 0.005$, * $p < 0.05$.

Dependent variable Sample of funds	Tracking Error			
	TE All	TE All	TE Single	TE Single
CRT	0.027 (0.547)	0.042 (0.315)	0.025 (0.628)	0.048 (0.345)
TOM	-0.034 (0.341)	-0.022 (0.564)	-0.032 (0.444)	-0.040 (0.313)
COMPETITIVENESS	0.071 (0.115)	0.090 (0.092)	-0.044 (0.351)	-0.051 (0.325)
RISK TOLERANCE	0.003 (0.486)	0.003 (0.494)	-0.003 (0.506)	-0.002 (0.775)
LOSS TOLERANCE	0.006 (0.322)	0.007 (0.226)	0.008 (0.169)	0.007 (0.257)
AMBIGUITY TOLERANCE	0.007* (0.023)		0.009** (0.002)	
AMBIGUITY NEUTRAL		0.097 (0.336)		0.235 (0.052)
AMBIGUITY SEEKING		0.228 (0.072)		0.354** (0.005)
$\min(R_{t-1}^{gross}, 0)$	-0.065** (< 0.001)	-0.065** (< 0.001)	-0.065** (< 0.001)	-0.066** (< 0.001)
$\max(R_{t-1}^{gross}, 0)$	0.015* (0.005)	0.015** (0.004)	0.011 (0.074)	0.011 (0.079)
$\log(AUM)_{t-1}$	-0.040* (0.012)	-0.035* (0.037)	-0.065** (0.003)	-0.062** (0.003)
EXPERIENCE	0.011* (0.041)	0.010 (0.063)	0.009 (0.129)	0.010 (0.080)
TEAM	-0.112 (0.092)	-0.126 (0.099)		
Constant	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	14,480	14,480	7,046	7,046
Number of managers	84	84	53	53
Adjusted R^2	0.588	0.574	0.642	0.636

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the wofo, answered on scales ranging from 1 to 7 each. $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month's gross return if it was negative; similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month's gross return if it was positive. $\log(AUM)$ stands for the log of assets under management, EXPERIENCE indicates years in industry, and TEAM is a dichotomous indicator for team-managed funds.

Robustness checks: We run several robustness checks. The corresponding results are tabulated in Appendix C.

First, we consider an alternative benchmark assignment. The first two steps of our original benchmark assignment procedure involve self-selected benchmarks: in the first step we use the fund’s prospectus benchmark (which is self-designated), while in the second step we use the most common prospectus benchmark among all funds from the same fund category (i.e., we also rely on self-designated benchmarks chosen by other funds). Sensoy (2009) argues that resulting from strategic behavior driven by the incentive to improve flows, almost one-third of actively managed US equity mutual funds specify a prospectus benchmark that does not match the fund’s actual investment style. To account for the concerns regarding strategic benchmark selection, we repeat the analysis using only the “Lipper Technical Indicator Benchmark”, which is independently assigned to each fund by Lipper. Note that 230 funds in the sample have a Lipper benchmark and these funds are managed by 80 of our managers (versus 267 funds managed by 84 managers for our original benchmark assignment). Also note that this procedure assigns a different benchmark to the majority of the funds, since only 46 funds are associated with the Lipper benchmark through our original assignment procedure. The first two columns of Table S5 show the results for relative volatility (compare to the first two columns of Table 4). The effect of cognitive reflection abilities becomes insignificant, but the magnitude of the point estimates remain qualitatively similar: e.g., in comparison to fund managers with low CRT scores, relative fund risk for fund managers with the highest possible CRT score is 7.5 lower (versus 9.5 lower in Table 4). The effect of risk preferences (the coefficient on RISK TOLERANCE and the related dummies) remains significant, and the magnitude of the effect even becomes slightly stronger. The last two columns of Table S5 show the results for tracking error (compare to the first two columns of Table 5). The coefficients on AMBIGUITY TOLERANCE and the related dummy variables remain very similar to those in our main specification both in terms of magnitude and statistical significance. Overall, the results seem generally robust to the alternative benchmark assignment that avoids self-designated benchmarks.

Second, we re-estimate our main specifications (i.e., using both single- and team-managed funds and using the raw experimental measures instead of corresponding dummies) in five-year sub-samples: 2008-2012 and 2013-2017. Similar to the results on the full sample, there is no significant relationship between the experimental measures and the fund performance measures (see Table S6). The results in Table S7 show that our findings on relative volatility and tracking error remain qualitatively robust, despite exhibiting partly higher standard errors resulting from lower statistical power due to the smaller sample sizes.

Third, Table S8 shows that the results are robust to including PATIENCE as an additional explanatory variable. As we pointed out previously, we lost a few observations on our experimental measure on inter-temporal preferences (10 participants) due to a runtime error issue in the experimental procedure. As is shown in Table S8, PATIENCE does not have a significant effect on any of the dependent variables we consider, and our conclusions on the effects of the other experimental measures remain unchanged when PATIENCE is included as an additional regressor. This is the reason why the variable PATIENCE is omitted from the main analysis.

The fourth robustness check is concerned with the period used for calculating the fund risk measures. In the main analysis we use daily returns over a month to calculate the relative volatility (RV_{it}) and tracking error (TE_{it}), and hence the resulting unit of observation is a fund-month. In this robustness check we use daily data over 6-month periods to calculate the risk measures, e.g.,

$$RV_{ih} = \frac{Std\left(R_{id[h]}^{net}\right)}{Std\left(R_{id[h]}^B\right)}, \quad (6)$$

where $R_{id[h]}$ denotes daily returns within half-year h (i.e., we have two observations within a year). The resulting unit of observation in these regressions is fund-half year. Columns 1 to 4 in Table S9 show that the results on RV and TE are robust to calculating these measures on different frequencies.

Finally, we consider an alternative measure of fund riskiness, referred to as relative semi-volatility, which is based on the target semi-deviations of returns with the target return being zero:

$$RSV_{ih} = \frac{Std\left(R_{id[h]}^{net} \mid R_{id[h]}^{net} < 0\right)}{Std\left(R_{id[h]}^B \mid R_{id[h]}^B < 0\right)}. \quad (7)$$

Longer periods than a month are needed to reliably estimate RSV to ensure that there is a suitable number of negative return observations, and therefore we only include RSV in the analysis when the risk measures are calculated on a half-year frequency. Comparing the first two and the last two columns of Table S9, we can see that the conclusions obtained when using relative semi-volatility as the dependent variable are very similar to those obtained when relative-volatility is used.

4. Conclusion

In this paper we address the question whether fund managers' cognitive skills, economic preferences, and attitudes toward competition can explain fund performance and dynamics. We recruited 84 fund managers, managing 267 mutual funds, from four large and mid-sized countries in the European Union. We matched the experimental data of the fund managers with the time series of the funds they manage and we controlled for various variables as, among others, the funds' benchmarks, fund categories, assets under management, and years of experience in all analyses.

First, we find a strong positive relationship between fund managers' risk tolerance and fund volatility. This indicates that fund managers with low (high) levels of risk tolerance, on average, compose funds with lower (higher) fund volatility, relative to the benchmark. Importantly, this finding holds while controlling for fund managers' self-selection into different fund categories as well as for other economic preferences and cognitive skills, and for additional variables like fund benchmark, fund category, industry experience, and fund size.

Second, we observe that fund managers with lower levels of ambiguity tolerance manage their funds with lower tracking errors compared to peers with higher levels of ambiguity tolerance. This indicates that ambiguity tolerance explains fund managers' propensity to "risk" deviating from the benchmark.

Finally, we report that cognitive skills do not explain fund performance in terms of neither abnormal returns nor value added. However, we do find that fund managers with high cognitive reflection abilities (CRT scores) generate these returns at lower risk.

Our results have several implications for the fund industry. First, they suggest that fund managers' cognitive skills matter for diversification and that economic preferences shape their allocation behavior. This is shown by those fund managers with high cognitive abilities generating a certain performance while taking less risk and by the fact that fund managers' risk tolerance positively correlates with fund risk. We consider these findings important, as monetary and non-monetary incentives (Kirchler et al., 2018b), the institutional framework in the industry (Chevalier and Ellison, 1999), and the competitive pressure of the market, that could limit or eliminate the role of professionals' preferences, do not eliminate their impact on fund allocation.

Second, from a labor market perspective, human resource managers could elicit cognitive skills and economic preferences in a more scientific way, achieving a higher probability of recruiting fund managers with superior skills and achieving a better match of fund risk and manager's risk preferences. However, at this stage we should be cautious in our interpretation as we cannot say much about persistence of our findings with our data set. We leave the answer to the question of persistence and causality for future research. In five to ten years we will analyze whether cognitive skills and economic preferences and attitudes, elicited in Winter 2017, were able to predict (risk-adjusted) abnormal returns, fund risk, and tracking errors.

Third, our results indicate that fund risk and risk-adjusted performance depend to a certain degree on the fund manager's preferences and abilities. From a customer's perspective, this has important consequences, as it is unclear whether these patterns fit the investors' preferences (who is not informed about the economic preferences and cognitive abilities of the fund manager). This observation is in line with the experimental finding of Kirchler et al. (2018a), showing that financial professionals' self-assessed risk attitude predominantly explains risk-taking on behalf of third parties (customers). Our findings relate to the empirical observations of Foerster et al. (2017) and Linnainmaa et al. (2019). Both studies show that advisor fixed effects explain considerably more variation in household portfolio risk than a broad set of investor attributes and that most financial advisors invest their personal portfolios just like they advise their clients—i.e., they trade too much, chase returns, prefer expensive, actively managed funds, and hold underdiversified portfolios.

Finally, from a scientific perspective, we interpret our study as a test for the predictive power of laboratory experiments regarding the universality of economic preferences. We contribute to the scientific debate by means of reporting an effect of experimentally elicited risk preferences on fund risk. Our results suggest that economic preferences elicited in experimental settings indeed have predictive power for real-world (risky) professional behavior.

References

- Abdellaoui, M., Bleichrodt, H., & Kammoun, H. (2013). Do financial professionals behave according to prospect theory? An experimental study. *Theory and Decision*, 74(3), 411–429.
- Adams, R., Keloharju, M., & Knüpfer, S. (2018). Are CEOs born leaders? lessons from traits of a million individuals. *Journal of Financial Economics*, 130, 392–408.
- Alevy, J. E., Haigh, M. S., & List, J. A. (2007). Information cascades: Evidence from a field experiment with financial market professionals. *The Journal of Finance*, 62(1), 151–180.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241–251.
- Berk, J. B., & van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, 118(1), 1–20.
- Bianchi, M. (2018). Financial literacy and portfolio dynamics. *The Journal of Finance*, 73(2), 831–859.
- Bianchi, M., & Tallon, J.-M. (2018). Ambiguity preferences and portfolio choices: Evidence from the field. *Management Science*, 65(84), 1455–1947.
- Bodnaruk, A., & Simonov, A. (2016). Loss-averse preferences, performance, and career success of institutional investors. *The Review of Financial Studies*, 29(11), 13–21.
- Böhm, M. J., Metzger, D., & Strömberg, P. (2018). *Since you’re so rich, you must be really smart: Talent and the finance wage premium*.
- Bosch-Rosa, C., Meissner, T., & Bosch-Domènech, A. (2018). Cognitive bubbles. *Experimental Economics*, 21, 132–153.
- Bruguier, A. J., Quartz, S. R., & Bossaerts, P. (2010). Exploring the nature of “trader intuition”. *The Journal of Finance*, 65(5), 1703–1723.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57–82.
- 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.
- Chevalier, J., & Ellison, G. (1999). Are some mutual fund managers better than others? cross-sectional patterns in behavior and performance. *The Journal of Finance*, 54(3), 875–899.
- Christelis, D., Jappelli, T., & Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review*, 54(1), 18–38.
- Cipriani, M., & Guarino, A. (2009). Herd behavior in financial markets: An experiment with financial market professionals. *Journal of the European Economic Association*, 7(1), 206–233.
- Corgnet, B., DeSantis, M., & Porter, D. (2015a). *Revisiting information aggregation in asset markets: Reflective learning and market efficiency*. Chapman University, Economic Science Institute.
- Corgnet, B., DeSantis, M., & Porter, D. (2018). What makes a good trader? on the role of intuition and reflection on trader performance. *The Journal of Finance*, 73(3), 1113–1137.
- Corgnet, B., Hernán-González, R., Kujal, P., & Porter, D. (2015b). The effect of earned versus house money on price bubble formation in experimental asset markets. *Review of Finance*, 19(4), 1455–1488.

- Cremers, M., & Petajisto, A. (2009). How active is your fund manager? a new measure that predicts performance. *Review of Financial Studies*, 22(9), 3329–3365.
- DeMartino, B., O’Doherty, J., Ray, D., Bossaerts, P., & Camerer, C. (2013). In the mind of the market: Theory of mind biases value computation during financial bubbles. *Neuron*, 79, 1222–1231.
- Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., & Peijnenburg, K. (2016). Prices in experimental asset markets under uncertainty. *Journal of Financial Economics*, 119, 559–577.
- Dorn, D., & Huberman, G. (2005). Talk and action: What individual investors say and what they do. *Review of Finance*, 9, 437–481.
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics*, 133(4), 1645–1692.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56.
- Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *The Journal of Finance*, 65(5), 1915–1947.
- Fang, Y., & Wang, H. (2014). *Fund manager characteristics and performance*.
- Foerster, S., Linnainmaa, J. T., Melzer, B. T., & Previtiero, A. (2017). Retail financial advice: Does one size fit all? *Journal of Finance*, 72(4), 1441–1482.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25–42.
- Gächter, S., Johnson, E. J., & Hermann, A. (2007). *Individual-level loss aversion in riskless and risky choices*. CeDEx Discussion Paper No. 2007-02.
- Golec, J. H. (1996). The effects of mutual fund managers’ characteristics on their portfolio performance, risk and fees. *Financial Services Review*, 5(2), 133–147.
- Graham, J. R., Harvey, C. R., & Puri, M. (2013). Managerial attitudes and corporate actions. *Journal of Financial Economics*, 109, 103–121.
- Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2011). IQ and stock market participation. *The Journal of Finance*, 66(6), 2121–2164.
- Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2012). IQ, trading behavior, and performance. *Journal of Financial Economics*, 104(2), 339–362.
- Grinblatt, M., & Titman, S. (1989). Mutual fund performance: An analysis of quarterly portfolio holdings. *The Journal of Business*, 62(3), 393–416.
- Grinblatt, M., & Titman, S. (1993). Performance measurement without benchmarks: An examination of mutual fund returns. *The Journal of Business*, 66(1), 47–68.
- Haigh, M. S., & List, J. A. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *The Journal of Finance*, 60(1), 523–534.
- Helmreich, R. L., & Spence, J. T. (1978). The work and family orientation questionnaire: An objective instrument to assess components of achievement motivation and attitudes toward family and career. *JSAS Catalog of Selected Documents in Psychology*, 8(35).
- Holzmeister, F. (2017). oTree: Ready-made apps for risk preference elicitation methods. *Journal of Behavioral and Experimental Finance*, 16, 33–38.

- Ibert, M., Kaniel, R., Van Nieuwerburgh, S., & Vestman, R. (2018). Are mutual fund managers paid for investment skill? *Review of Financial Studies*, 31(2), 715–772.
- Kahneman, D. (2011). *Thinking, fast and slow*. Ney York: Farrar, Straus and Giroux.
- Kaplan, S. N., Klebanov, M. M., & Sorensen, M. (2012). Which ceo characteristics and abilities matter? *The Journal of Finance*, 67(3), 973–1007.
- Kaustia, M., Alho, E., & Puttonen, V. (2008). How much does expertise reduce behavioral biases? The case of anchoring effects in stock return estimates. *Financial Management*, 37(3), 391–412.
- Kirchler, M., Lindner, F., & Weitzel, U. (2018a). *Delegated decision making and social competition in the finance industry*.
- Kirchler, M., Lindner, F., & Weitzel, U. (2018b). Rankings and risk-taking in the finance industry. *The Journal of Finance*, 73(5), 2271–2302.
- Linnainmaa, J. T., Melzer, B., & Previtro, A. (2019). The misguided beliefs of financial advisors. *The Journal of Finance*, forthcoming.
- Mackintosh, N. (2011). History of theories and measurement of intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *The cambridge handbook of intelligence* (pp. 3–19). Cambridge University Press.
- Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *The Journal of Finance*, 60, 2661–2700.
- Meier, S., & Sprenger, C. D. (2015). Temporal stability of preferences. *The Review of Economics and Statistics*, 97(2), 273–286.
- Noussair, C. N., Tucker, S., & Xu, Y. (2016). Futures markets, cognitive ability, and mispricing in experimental asset markets. *Journal of Economic Behavior & Organization*, 130, 166–179.
- Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, 72(1), 147–152.
- Primi, C., Morsanyi, K., Donati, M. A., & Hamilton, J. (2015). The development and testing of a new version of the cognitive reflection test applying item response theory (irt). *Journal of Behavioral Decision Making*, 29(5), 453–469.
- Raven, J. (2000). The Raven’s progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, 41(1), 1–48.
- Riedl, A., & Smeets, P. (2017). Why do investors hold socially responsible mutual funds? *The Journal of Finance*, 72(6), 2505–2550.
- Sensoy, B. A. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics*, 92(1), 25–39.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39(7), 1275–1289.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning*, 20(2), 147–168.
- Von Gaudecker, H.-M. (2015). How does household portfolio diversification vary with financial literacy and financial advice? *Journal of Finance*, 70, 489–507.
- Wermers, R. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *The Journal of Finance*, 55(4), 1655–1703.

Appendix

A. Details on the Experiment

In this section, the experimental tasks, the feedback map for the fund managers, and results of the various experimental tasks are described in more detail. The experiment has been conducted online using *oTree* (Chen et al., 2016). The software, including all instructions as used for the data collection, is available for download as a zipped *oTree* project at <https://osf.io/dq3t8/> and as a live demo version via <https://fea-2018-en.herokuapp.com>. Participants have been invited via hard-copy letters and/ or e-mail, based on contact information available via funds' fact sheets, the webpages of institutions, and *Morningstar*.

A.1. Feedback Map

At the end of the experiment, fund managers could indicate whether they wished to receive personalized feedback (as a multipage *.pdf-file distributed via e-mail) once the data collection has been completed. The feedback maps contained general information about each task and why the measured skill may potentially matter for financial decision-making. Moreover, subjects received their own scores as well as summary statistics about the performance of their peers participating in our experiment. Figure S1 show the title page and, as one example, the feedback pages for Raven's Advanced Progressive Matrices (APM).¹⁷

¹⁷ Note that the sample feedback shown in Figure S1 includes the full sample of 94 participants who completed the experiment. Since no fund data could be obtained for 10 participants, the remainder of this section refers to the sample of $n = 84$ fund managers, as do all analyses presented in the main text.

Feedback

ON YOUR PERSONAL AND YOUR PEERS' PERFORMANCE IN AN ONLINE EXPERIMENT WITH FINANCE PROFESSIONALS

Participant ID: 2576

Centre for Finance, University of Gothenburg, Sweden
Department of Banking and Finance, University of Innsbruck, Austria

Fluid Intelligence: Raven's Advanced Progressive Matrices

What is it about?

Raven's Advanced Progressive Matrices (often referred to simply as Raven's Matrices) is a non-verbal test typically used in educational settings.

It measures abstract reasoning and is regarded as a non-verbal estimate of fluid intelligence. Fluid intelligence or fluid reasoning is the capacity to reason and solve problems independent of any knowledge from the past.

The Raven's Advanced Progressive Matrices test is the most common and popular test administered to groups ranging from 5-year-olds to the elderly.

Why does it matter?

Fluid intelligence describes the ability to analyze novel problems and identify patterns and relationships that underpin these problems. This skill is necessary for all logical problem solving in scientific, mathematical, and technical questions – tasks that very often arise in financial decision making.

How did you and your peers score?

The graph below depicts the distribution of correct answers to the 18 matrices in this experiment (number of observations N = 94). On average, 10 questions were answered correctly (median). Moreover, those in the top quartile solved at least 12 questions correctly while those in the bottom quartile solved up to 8 questions correctly.

The number of questions you have answered correctly is highlighted in blue so that you can compare yourself to the sample. Note that for time reasons we have only selected 18 representative questions from the entire questionnaire which makes comparisons with other samples impossible.

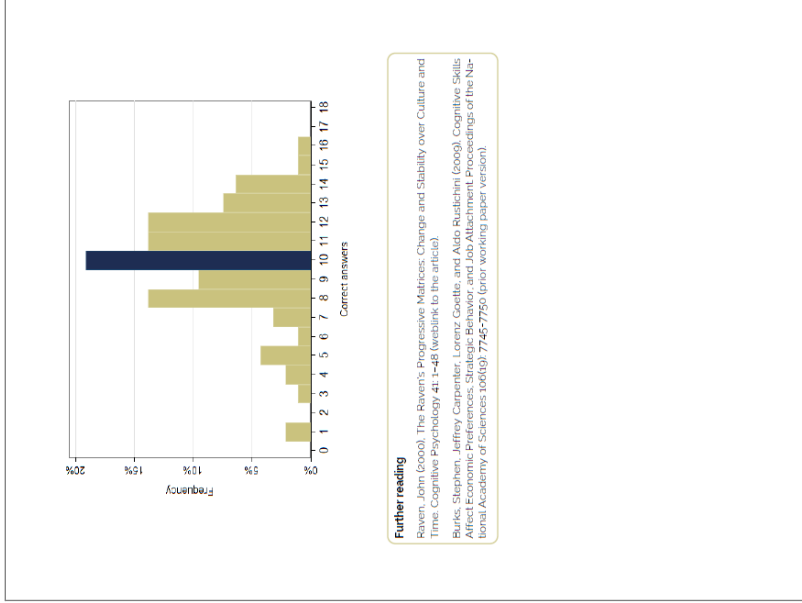


Figure S1: Feedback map: The figure shows the title page (left) and, as one example, the information and feedback provided for the Advanced Progressive Matrices task (APM; middle and right). Background information and feedback comparing the individual performance to all participating peers has been presented in a similar way for all other measures elicited in the experiment.

A.2. Details on the Experimental Tasks

Below, we provide further details on the experimental protocols of the tasks used to elicit cognitive abilities and economic preferences in the online experiment. The distributions of scores in the experimental tasks are depicted in Figure S3; correlations between the measures are summarized in Table S1.

Fluid Intelligence. Raven’s Advanced Progressive Matrices (APM; Raven, 2000) are designed to measure fluid intelligence. We presented subjects with 18 increasingly difficult items (instead of the 36 items in the original version) where they had to infer the missing element of a given diagrammatic puzzle. In particular, we used the every second item, starting with the first puzzle. For further details, we refer to the demo version of the software (<https://fea-2018-en.herokuapp.com>).

Cognitive Reflection Test. Cognitive reflection tests are designed to measure subjects’ ability to consciously reflect on their intuitive responses. These types of tests were first established by Frederick (2005) and have been used widely since. To avoid potential recognition effects by the subjects, we decided to use questions from newer versions of the test proposed by Toplak et al. (2014) and Primi et al. (2015). Each question was displayed on a separate screen; the order has been randomized to avoid order effects. In particular, we included the following five questions (correct answers in parentheses):

- If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together? (4 days)
- Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class? (29 students)
- A man buys a pig for \$60, sells it for \$70, buys it back for \$80, and sells it finally for \$90. How much has he made? (\$20)
- If three elves can wrap three toys in one hour, how many elves are needed to wrap six toys in two hours? (3 elves)
- In an athletics team, tall members are three times more likely to win a medal than short members. This year the team has won 60 medals so far. How many of these have been won by short athletes? (15 medals)

Theory of Mind. Theory of mind is a social sensitivity-skill that refers to the capacity of “reading the minds” of other people. In the “Reading-the-Mind-in-the-Eyes”-test introduced by Baron-Cohen et al. (2001), subjects have to infer the emotional state of a person from a picture showing only their eye region. In each of 18 trials, subjects had to select the correct emotion from a list of four adjectives. For each of the four potential answers, we provided participants with synonyms and an example sentence using the adjective (describing an emotional state) in an easy-to-understand context. To make sure participants understand the task, we implemented one practice trial, providing them with feedback about whether their choice has been correct. We used the same subset of 18 pictures as Corgnet et al. (2018). The number of correctly chosen emotions serves as our measure of “reading-the-mind”-skills (TOM). For further details, we refer to the demo version of the software (<https://fea-2018-en.herokuapp.com>).

Risk Preferences. The staircase risk elicitation method by Falk et al. (2018) allows to infer a subject's certainty equivalent for a given lottery, ensuring consistent answers. In four iterative, path-dependent questions, subjects decided between a lottery that pays €60 or €0 with equal probability and a certain payment that varies from question to question (see Figure S2 for a graphical representation of the task). Precisely, due to the limited number of iterations, the staircase approach yields intervals for the certainty equivalents. The midpoints of the intervals constitute our measure for subjects risk attitudes (RISK TOLERANCE). The task was implemented using the ICL app put forward by Holzmeister (2017).

Time Preferences. We used the same staircase approach proposed by Falk et al. (2018) to measure time preferences, implemented via a modified version of the ICL app of Holzmeister (2017). In four path-dependent questions, subjects had to decide whether they preferred a payment of €20 today or a certain higher amount in 6 months. The future premium increased in the next question when a subject opted for the payment today and decreased when the subject chose the future payment. Similarly to the risk elicitation task, this approach yields intervals for the time premia required by subjects to wait 6 months. The intervals' midpoints serve as our measure for participants' patience (PATIENCE). For a facilitated interpretation, we compute the future premium subjects were willing to give up in order to receive the payment today (i.e. we multiplied the time premia by -1). Thus, higher values represent higher patience.

Loss Tolerance. The task to elicit participants' attitudes toward losses is based on the exercise proposed by Gächter et al. (2007). However, to align the task with the experiments to elicit risk and time preferences, we transformed the elicitation procedure into an interactive, path-dependent series of questions, utilizing the ICL app put forward by Holzmeister (2017). In each question, subjects decided whether they wished to participate in a lottery paying either €22 or some negative amount with equal probability. In the end, the task reveals intervals for each subject's maximum accepted loss in order to have the chance to win €22 (LOSS TOLERANCE).

Ambiguity Tolerance. We followed the setup of Dimmock et al. (2016) to elicit ambiguity tolerance. As in the original task, subjects had to choose between two urns containing 100 balls of blue and orange color. At the end of the experiment, one ball was drawn randomly from the chosen box. If the ball was blue, the subject would win €60; if the ball was orange, the subject would win nothing. While the distribution of blue and orange balls (i.e., the probability of winning) was known in the first box, the probability was unknown for the second box. In the first decision, the known distribution offered a 50% chance of winning. Subjects who chose the known distribution (risk) were presented a lower known probability of winning in the second question, while subject who chose the unknown distribution (ambiguity) were presented a higher known probability of winning in the second question. This procedure reveals a matching probability that leaves subjects indifferent between the risky and the ambiguous alternative, constituting our measure (AMBIGUITY TOLERANCE) of ambiguity preferences.

Competitiveness. The Work and Family Orientation (wofo) questionnaire of Helmreich and Spence (1978) is a widely used psychometric measure of individuals' competitiveness. Subjects answered how strongly they agree with a certain statement about their attitudes towards competition on a scale from 1 to 7. The competitiveness score (COMPETITIVENESS) is then computed as the sum of the individual answers. In particular, participants answered the following five questions:

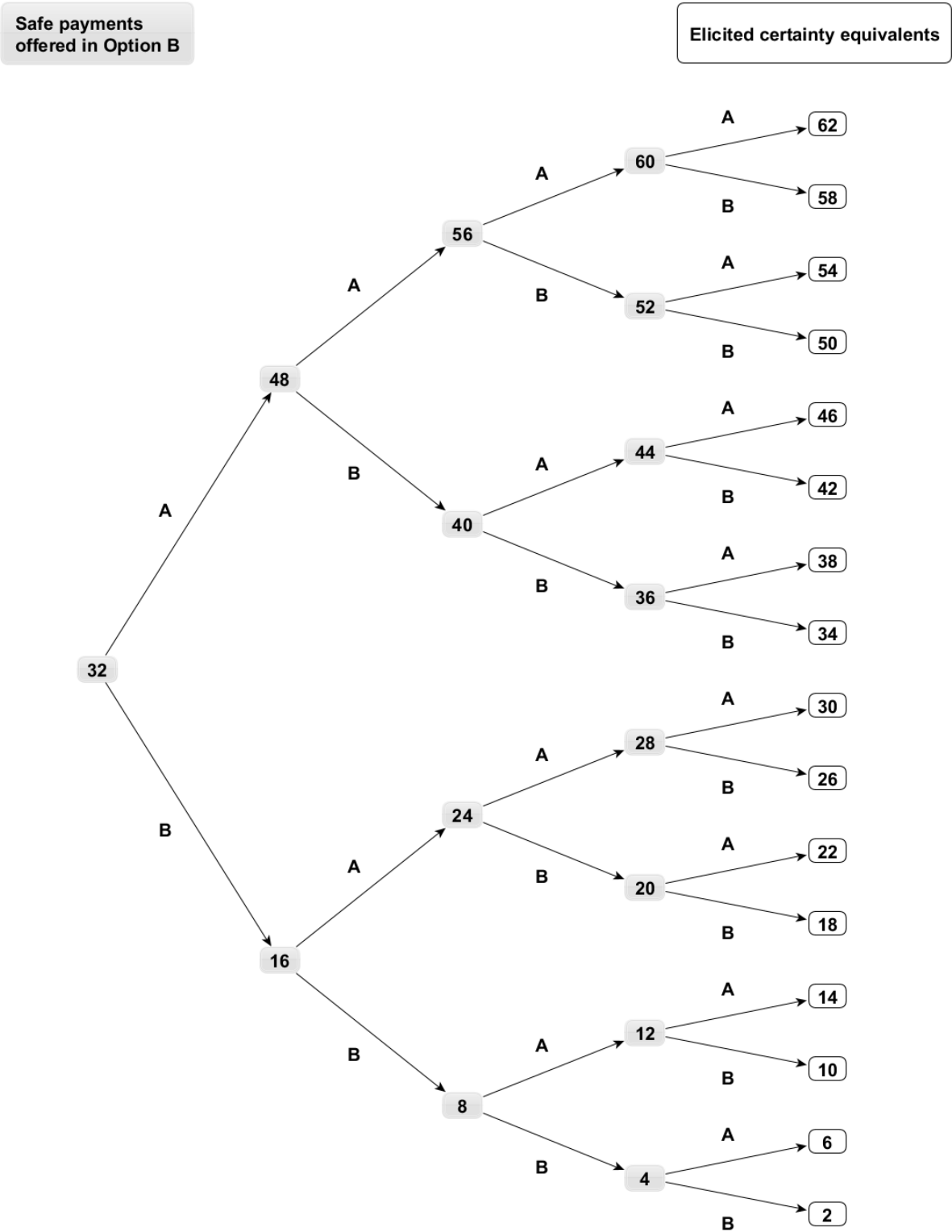


Figure S2: The staircase risk preference elicitation procedure is based on Falk et al. (2018). The final column shows the midpoints of the elicited intervals for the certainty equivalents (± 2 Euros). The iterative methods for eliciting attitudes towards time discounting, losses, and ambiguous outcomes have been implemented based on the same structuring.

- I enjoy working in situations involving competition with others.
- It is important to me to perform better than others on a task.
- I feel that winning is important in both work and games.
- It annoys me when other people perform better than I do.
- I try harder when I'm in competition with other people.

A.2.1. Correlation Between the Experimental Variables

Table S1: Pearson correlation coefficients between the experimental variables. ** $p < 0.005$, * $p < 0.05$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) CRT	1.00						
(2) APM	0.52**	1.00					
(3) TOM	0.11	0.01	1.00				
(4) RISK TOLERANCE	0.15	0.13	0.02	1.00			
(5) LOSS TOLERANCE	-0.07	0.01	-0.12	0.31**	1.00		
(6) AMBIGUITY TOLERANCE	0.09	0.25*	0.05	0.09	0.03	1.00	
(7) COMPETITIVENESS	0.20	0.17	0.05	-0.12	-0.07	0.10	1.00

Note: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the wofo, answered on scales ranging from 1 to 7 each.

B. Details on the Fund Data

B.1. Matching Funds to Participants

We start with the 94 fund managers who participated in our experiments, and use the *Morningstar Direct* database to match the managers to their funds. We obtain time series data on all open-ended mutual funds (both active and inactive) that have one of our four countries registered as "Domicile" or "Region of Sale" in the database. *Morningstar* provides the manager history for each fund, which contains the full name of the fund's managers (current and past) together with the tenure for each manager (start date and end date).

In a first step, we searched for all the funds (i) where the full name of one of our participants appears as the fund's manager during the period from January 2008 to December 2017, and (ii) where the fund company matches the participant's workplace at the time of the experiment. We manage to match at least one fund to 93 of our 94 participants and identify 378 funds altogether.¹⁸ In a second step, we

¹⁸ For one participant, we are not able to match any of the funds. According to this participant's LinkedIn profile, she/he started to work for the given company in June 2016, so she/he might have been too junior to be listed as a manager for any particular fund at the time of our experiment.

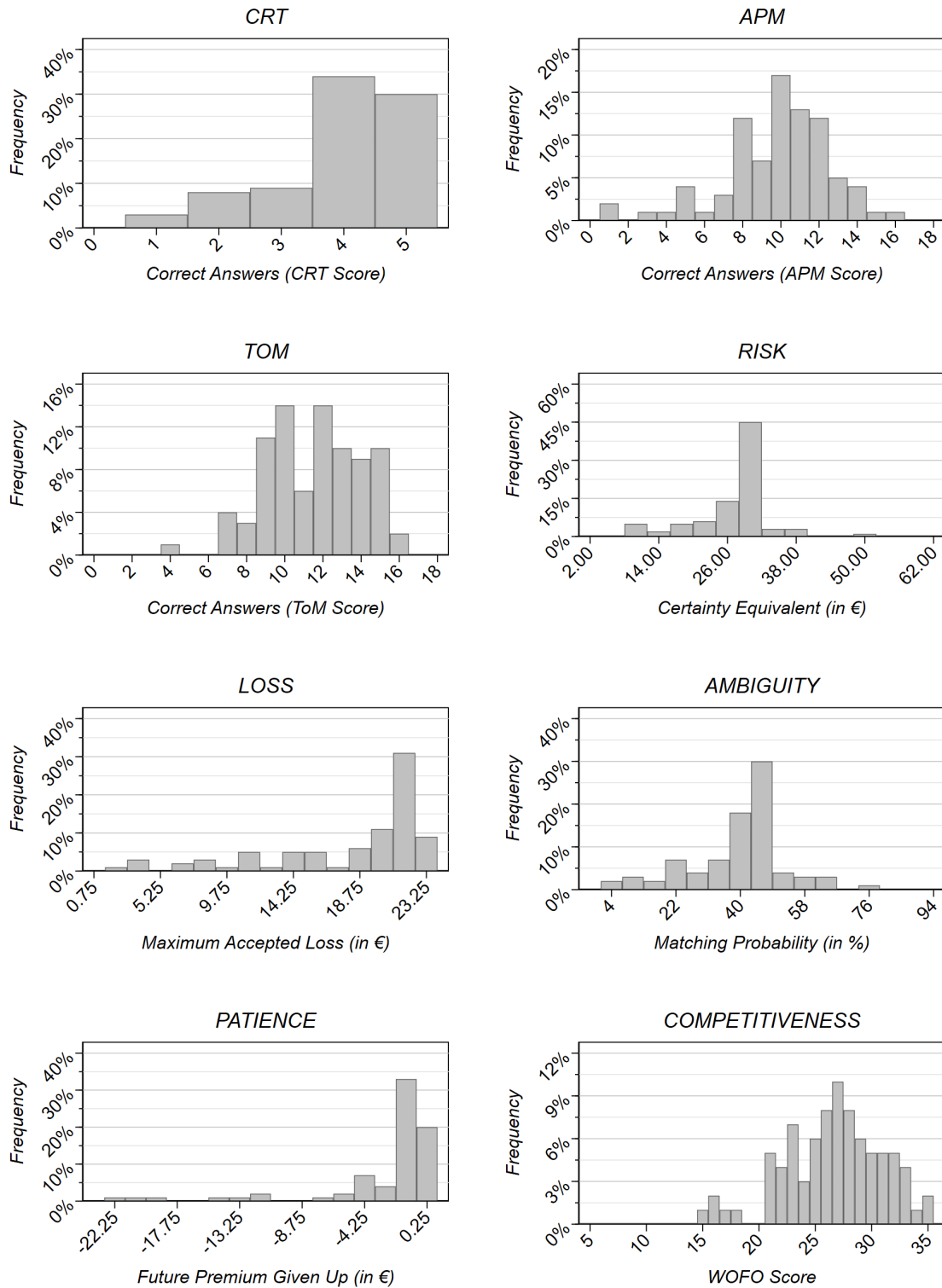


Figure S3: Histograms of the eight tasks on cognitive skills, economic preferences, and attitudes towards competitiveness. The values for economic preferences represent midpoints of the corresponding elicited intervals. The scaling of the x -axes ranges from the potential minimum to the potential maximum value of the task.

Table S2: Number of (non-missing) fund-month observations for each variable.

	Fund-month observations	Number of Funds
Sample	20,734	369
Assets under management (<i>AUM</i>)	19,694	364
Net return (in €)	20,259	369
Total Expense Ratio (<i>ER</i>)	19,953	362

augment the compiled list with funds where the full name of one of our participants appears as manager during the sample period, but the fund company does not match the participant’s workplace at the time of the experiment if two conditions are satisfied: (i) the manager’s tenure at this fund ended before December 2017, and (ii) we are able to verify that it is likely to be the same person (e.g., through the manager’s biography on the current employer’s website or the manager’s LinkedIn profile). This results in 14 additional funds. In a third step, following the literature (see, e.g., Ibert et al., 2018), we eliminate money market mutual funds (7 funds identified as such by the variable “Broad Category Group” in *Morningstar*) and index funds (9 funds identified as such by *Morningstar* or by the word “index” in their name). Finally, we exclude 7 funds for which we are not able to obtain any fund-level time-series data (the details of collecting the fund-level data is described in subsection B.2). Our resulting base sample consists of 369 funds managed by 93 managers for a total of 20,734 unique fund-month observations.¹⁹ Note that we do not lose any managers when excluding the money market funds, index funds, and funds with no time-series data.

To ensure anonymity, after matching the participants to their funds, the managers’ names are replaced with randomly generated, unique identifiers to match the de-personalized fund data with the experimental data at a later stage.

B.2. Monthly and Daily Fund-Level Data

In this section we describe how the fund-level data on assets under management, returns, and expense ratios are obtained. Mutual funds in our sample can have multiple share classes. We start with data on the share classes (identified by ISIN code) and then aggregate to the fund level. Table S2 provides a brief description of the number of non-missing fund-month observations for each variable. In the following subsections, we provide details for the three variables of interest for our research questions.

B.2.1. Assets Under Management

We start by assembling share class level assets under management (*AUM*) data using the following steps:

- **Step 1:** We retrieve share class level data from *Morningstar*.
- **Step 2:** We also retrieve share class level data from *Lipper*. For share classes with no *AUM* data at all in *Morningstar* (i.e., the whole *AUM* series is missing), we use *AUM* values from *Lipper*, if available.

¹⁹ For a given manager-fund observation, if the start date (end date) does not coincide with the first (last) day of the month, we include that month in the manager’s tenure if she/he was the manager for at least 20 calendar days during that month.

Table S3: Summary of steps involved for creating fund level observations for *AUM*.

		Frequency	Percent
Step 1:	Only <i>Morningstar</i> data	18,493	93.90%
Step 2:	Only <i>Lipper</i> data	111	0.56%
Step 3:	<i>Morningstar</i> and <i>Lipper</i> data	977	4.96%
Step 4:	Imputations needed	113	0.58%

- **Step 3:** For share classes where we have *AUM* data from both *Morningstar* and *Lipper* we calculate

$$AUMratio_{it} = \frac{AUM_{it}^{Morningstar}}{AUM_{it}^{Lipper}} \quad (8)$$

for months where both datasets report an observation. Then we replace the missing *Morningstar* values with *Lipper* values for the share classes where $0.99 < \overline{AUMratio}_i < 1.01$, i.e., where the *AUM* values from the two datasets are very close to each other in those months when both datasets report a value.

- **Step 4:** We impute the missing *AUM* values following Ibert et al., 2018. Only missing values in the middle of *AUM* series are imputed using past *AUM* values, the return on the share class, and a factor that adjusts for flow rates. Let $[t_0, t]$ and $[t + n, T]$ be periods where the share class has *AUM* data, i.e., $[t + 1, t + n - 1]$ is the period with missing values. The missing values are then filled according to

$$AUM_{ik} = F_i \cdot AUM_{ik-1} (1 + R_{ik}) \quad \text{for } k \in [t + 1, t + n - 1] \quad (9)$$

$$F_i = \left(\frac{1}{\prod_{k=t+1}^{t+n} (1 + R_{ik})} \cdot \frac{AUM_{it+n}}{AUM_{it}} \right)^{\frac{1}{n}}, \quad (10)$$

where i denotes the share class, F_i is the factor adjusting for flow rate during the missing period, and R_{ik} is the net return on the share class in month k .

After assembling the share class level *AUM* data, we aggregate it to the fund level. That is, for each fund-month observation, we aggregate *AUM* across all share classes of the fund to get the fund-level *AUM*. Altogether, we have 19,694 non-missing fund-month *AUM* observations. Table S3 summarizes which of the above steps are involved when creating the fund-level observations. The table reveals that the majority of the fund-month observations (93.9%) rely solely on the *Morningstar* database.

B.2.2. Net Return

Monthly frequency. We retrieve share class level monthly net returns from *Morningstar*. All returns are converted into Euro for comparability across different countries. *Morningstar* has a very thorough coverage of returns, so we do not augment the data with additional sources. We aggregate the share class level returns to the fund level. Altogether, we have 20,259 non-missing fund-month return observations. When aggregating to the fund level, we have the following options:

- If the fund has only one share class in the given month, we take the return on the single share class (54.1% of the observations).

- If the fund has multiple share classes in the given month but we observe the same return across all share classes with a return observation, we take this common return (9.0% of the observations).
- If the fund has multiple share classes in the given month and the return varies across the share classes, we take the *AUM*-weighted return across the share classes (36.8% of the observations).
- If the fund has multiple share classes in the given month and we have return observations on multiple share classes but no *AUM* observations on any of the share classes, we take the simple average of the return across the share classes (0.1% of the observations).

Daily frequency. In order to measure the riskiness of mutual funds, we rely on daily fund returns. Therefore, we also retrieve share class level daily net returns from *Morningstar* for the relevant funds over the sample period. All returns are converted into Euro for comparability across different countries. Then we aggregate the returns to the fund-level, yielding a total of 455,597 non-missing fund-day return observations. When aggregating to the fund level, similar to the case of the monthly returns, we have the following options:

- If the fund has only one share class in the given month, we take the return on the single share class (54.1% of the observations).
- If the fund has multiple share classes in the given month but we observe the same return across all share classes with a return observation, we take this common return (5.7% of the observations).
- If the fund has multiple share classes in the given month and the return varies across the share classes, we take the *AUM*-weighted return across the share classes (40.1% of the observations).
- If the fund has multiple share classes in the given month and we have return observations on multiple share classes but no *AUM* observations on any of the share classes, we take the simple average of the return across the share classes (0.1% of the observations).

When *AUM* data is needed, we use the monthly *AUM* described in Section B.2.1.

B.2.3. Total Expense Ratio

Expense ratios (*ER*) are reported yearly in both *Morningstar* and *Lipper*, so we start by assembling yearly share class level *ER* data using the following steps:

- **Step 1:** We retrieve share class level data from *Morningstar*.
- **Step 2:** We also retrieve share class level data from *Lipper*. For share classes with no *ER* data at all in *Morningstar* (i.e., the whole *ER* series is missing), we use *ER* values from *Lipper*, if available.
- **Step 3:** For share classes where we have *ER* data from both *Morningstar* and *Lipper* we calculate

$$ERdiff_{it} = \left| ER_{it}^{Morningstar} - ER_{it}^{Lipper} \right|, \quad (11)$$

for years where both datasets have an observation. Then we replace the missing *Morningstar* values with *Lipper* values for the share classes where $\overline{ERdiff}_i < 10bps$ and $\max(ERdiff_{it}) < 25bps$, i.e., where the mean difference is not larger than 10 basis points per year and the maximum difference is not larger than 25 basis points per year during those years when they both datasets report a value.

- **Step 4:** For share classes where the ER series from *Morningstar* stops earlier than the ER series from *Lipper*, and the two databases had exactly the same value for the last available joint observation, we use the *Lipper* values for the remaining period.

After the steps outlined above, we use imputations similar to Ibert et al., 2018:

- **Step 5:** For share classes where the ER series is constant, i.e., the smallest ER is equal to the largest ER for all existing observations, the missing ER observations are filled with this constant value.
- **Step 6:** For share classes that have missing values in the middle of ER series, the missing values are imputed using past ER values and ER growth rates. Let $[t_0, t]$ and $[t+n, T]$ be periods where the share class has ER data, i.e., $[t+1, t+n-1]$ is the period with missing values. The missing values are then filled according to

$$ER_{ik} = \left(\frac{ER_{it+n}}{ER_{it}} \right)^{\frac{1}{n}} \cdot ER_{ik-1} \quad \text{for } k \in [t+1, t+n-1], \quad (12)$$

where i denotes the share class.

- **Step 7:** For share classes that have missing values at the tails of the ER series, we test if the ER series follow a linear time trend. If they do, we replace the missing ER values with the forecast values from the linear model. Let $[t_0, t]$ and $[t+n, T]$ be periods where the share class has missing ER observations, i.e., $[t+1, t+n-1]$ is the period with ER data. We estimate the model

$$\log(ER_{ik}) = a_i + b_i k + \varepsilon_{ik} \quad \text{for } k \in [t+1, t+n-1], \quad (13)$$

and fill the missing ER values

$$ER_{ik} = \exp(\hat{a}_i + \hat{b}_i k) \quad \text{for } k \in [t_0, t] \cup [t+n, T]. \quad (14)$$

if the p -value of \hat{b}_i is less than or equal to 5% and $n \geq 6$. If these conditions are violated, we fill the missing ER values at the left (right) tail of the series with the mean values of the first (last) three ER values.

After the above steps, we create monthly share class level data by assigning $\frac{ER_{it}}{12}$ to each month of year t for share class i . Then we aggregate the share class-level expense ratios to the fund level. Altogether, we have 19,953 non-missing monthly ER observations. When aggregating to the fund level, we have the following options:

- If the fund has only one share class in the given month, we take the ER on the single share class (45.8% of the observations).
- If the fund has multiple share classes in the given month but we observe the same ER across all share classes with a ER observation, we take this common ER (26.5% of the observations).
- If the fund has multiple share classes in the given month and ER varies across the share classes, we take the AUM -weighted ER across the share classes (26.9% of the observations).
- If the fund has multiple share classes in the given month and we have ER observations on multiple share classes but no AUM observations on any of the share classes, we take the simple average of ER across the share classes (0.9% of the observations).

Table S4: Summary of steps involved for creating fund level observations for *ER*.

		Frequency	Percent
Step 1:	Only <i>Morningstar</i> data	7,883	39.51%
Step 2:	Only <i>Lipper</i> data	918	4.60%
Step 3–4:	<i>Morningstar</i> and <i>Lipper</i> data	5,634	28.24%
Step 5–7:	Imputations needed	5,518	27.65%

Table S4 summarizes how much we rely on the different data sources and imputations when creating the fund level observations. The table reveals that we have to rely more heavily on *Lipper* and imputations to gather the *ER* data, than for returns and *AUM*.

B.3. Benchmark Assignment

Prospectus benchmark. A prospectus benchmark is reported in *Morningstar* for 249 funds in our sample (64% of all funds). Some funds have linear combinations of indices as their benchmark (62 of the 249). We find monthly returns for the benchmark indices, expressed in Euro, on *Morningstar*, *Lipper*, or *Datastream*. Then the returns on the benchmarks are assigned to the respective funds; when the benchmark is a linear combination of different indices, the benchmark return is only calculated for those fund-months when return data on all benchmark constituents are available. Altogether, we are able to assign a prospectus benchmark return to 10,077 fund-month observations, covering 199 funds (80% of the 249 that have a reported prospectus benchmark and 51% of our full sample).

Lipper benchmark. *Lipper* independently assigns the “Lipper Technical Indicator Benchmark” to most of the funds in the database according to its assessment of the fund’s investment strategy. The technical indicator benchmark is assigned to 358 funds in our sample (91% of all funds). We find monthly returns for the benchmark indices, expressed in Euro, on *Morningstar*, *Lipper*, or *Datastream*. Altogether, we are able to assign a *Lipper* benchmark return to 13,239 fund-month observations, covering 253 funds (71% of the 358 that have an assigned *Lipper* benchmark and 65% of our full sample).

Hand assigned benchmark. We also assign benchmarks “by hand” to equity funds in our sample (identified as such by the variable “Broad Category Group” in *Morningstar*). For each equity fund category defined by the *Morningstar* variable “Category”, we find the most common benchmark among all open-ended mutual funds (both active and inactive) that have one of our four countries registered as “Domicile”. This most common benchmark is assigned to all the funds in the given category. Altogether, we have a “manually” assigned benchmark return for 6,885 fund-month observations, covering 132 funds.

C. Robustness Checks

Table S5: Robustness check for relative volatility and tracking error using only the Lipper Technical Indicator Benchmark: The table shows the results of ordinary least squares regressions of funds' relative volatility and tracking error on cognitive skills and economic preferences/attitudes. The relative volatility, RV , measures the overall riskiness of the fund relative to the riskiness of the benchmark. The tracking error, TE , measures fund's net return minus the benchmark return. For each fund, we use the benchmark independently assigned by Lipper. Standard errors are clustered at the manager level. Corresponding p -values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

<i>Dependent variable</i>	<i>RV</i>	<i>RV</i>	<i>TE</i>	<i>TE</i>
<i>Sample of funds</i>	<i>All</i>	<i>All</i>	<i>All</i>	<i>All</i>
CRT	-0.019 (0.235)		0.012 (0.832)	0.023 (0.678)
CRT = 4		0.025 (0.503)		
CRT = 5		-0.075 (0.069)		
TOM	-0.009 (0.578)	-0.004 (0.787)	-0.041 (0.314)	-0.026 (0.516)
COMPETITIVENESS	0.009 (0.486)	0.009 (0.528)	0.086 (0.068)	0.106 (0.056)
RISK TOLERANCE	0.007** (< 0.001)		0.003 (0.621)	0.003 (0.581)
RISK NEUTRAL		0.090* (0.006)		
RISK SEEKING		0.168** (< 0.001)		
LOSS TOLERANCE	0.001 (0.768)	0.001 (0.630)	0.003 (0.695)	0.005 (0.544)
AMBIGUITY TOLERANCE	-0.001 (0.288)	0.000 (0.650)	0.007* (0.013)	
AMBIGUITY NEUTRAL				0.108 (0.348)
AMBIGUITY SEEKING				0.220 (0.097)
Constant & Controls	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	12,189	12,189	12,383	12,383
Number of managers	80	80	80	80
Adjusted R^2	0.170	0.193	0.575	0.560

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the wofo, answered on scales ranging from 1 to 7 each. Controls refer to the nonlinear transformations of the past gross return ($\min(R_{t-1}^{gross}, 0)$ and $\max(R_{t-1}^{gross}, 0)$), lagged log of assets under management ($\log(AUM)_{t-1}$), years in industry (EXPERIENCE), and a dichotomous indicator for team-managed funds (TEAM). In order to save space, the coefficients on the control variables are not reported.

Table S6: Robustness check for performance variables for the 5-year sub-periods 2008–2012 and 2013–2017, respectively: The table shows the results of ordinary least squares regressions of funds’ abnormal returns and values added on cognitive skills and economic preferences/attitudes. The abnormal return, R^{abn} , is the total monthly return of the fund over its benchmark. The value added, V , is the product of assets under management, and abnormal return. Standard errors are clustered at the manager level. Corresponding p -values are shown in parentheses. ** $p < 0.005$, * $p < 0.05$.

<i>Dependent variable</i>	R^{abn}	R^{abn}	V	V
<i>Time period</i>	08-12	13-17	08-12	13-17
CRT	−0.017 (0.598)	−0.018 (0.512)	0.014 (0.781)	0.044 (0.550)
TOM	−0.028 (0.586)	−0.010 (0.636)	−0.017 (0.746)	−0.014 (0.786)
COMPETITIVENESS	−0.011 (0.842)	0.038 (0.184)	0.065 (0.444)	0.061 (0.410)
RISK TOLERANCE	−0.001 (0.863)	−0.001 (0.766)	0.007 (0.503)	0.008 (0.535)
LOSS TOLERANCE	−0.013 (0.142)	0.003 (0.482)	−0.006 (0.441)	−0.003 (0.749)
AMBIGUITY TOLERANCE	0.003 (0.240)	0.003 (0.087)	0.002 (0.469)	0.004 (0.241)
$\log(AUM)_{t-1}$	0.023 (0.398)	0.016 (0.176)		
EXPERIENCE	−0.007 (0.269)	0.004 (0.092)	0.001 (0.937)	0.001 (0.921)
TEAM	0.104 (0.216)	0.092* (0.020)	0.261 (0.060)	0.274* (0.027)
Constant	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	4,573	10,052	4,573	10,052
Number of managers	58	84	58	84
Adjusted R^2	0.174	0.083	0.107	0.063

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the “Reading-the-Mind-in-the-Eyes”-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the wofo, answered on scales ranging from 1 to 7 each. $\log(AUM)$ stands for the log of assets under management, EXPERIENCE indicates years in industry, and TEAM is a dichotomous indicator for team-managed funds.

Table S7: Robustness check for relative volatility, and tracking error for the 5 years sub-periods 2008–2012 and 2013–2017, respectively: The table shows the results of ordinary least squares regressions of funds’ relative volatility and tracking error on cognitive skills and economic preferences/attitudes. The relative volatility, RV , stands for the overall riskiness of the fund relative to the riskiness of the benchmark. The tracking error, TE , measures fund’s net return minus the benchmark return. Standard errors are clustered at the manager level. Corresponding p -values are shown in parentheses. ** $p < 0.005$, * $p < 0.05$.

<i>Dependent variable</i> <i>Time period</i>	<i>RV</i>	<i>RV</i>	<i>TE</i>	<i>TE</i>
	08-12	13-17	08-12	13-17
CRT	−0.033 (0.053)	−0.033* (0.026)	0.024 (0.730)	0.029 (0.479)
TOM	−0.015 (0.605)	−0.008 (0.706)	−0.051 (0.394)	−0.022 (0.513)
COMPETITIVENESS	−0.001 (0.977)	0.020 (0.167)	0.043 (0.489)	0.095* (0.029)
RISK TOLERANCE	0.008** (< 0.001)	0.004* (0.013)	0.006 (0.324)	0.001 (0.847)
LOSS TOLERANCE	−0.003 (0.463)	0.002 (0.621)	0.008 (0.430)	0.005 (0.263)
AMBIGUITY TOLERANCE	−0.001 (0.395)	−0.002 (0.139)	0.009* (0.012)	0.005 (0.077)
$\min(R_{t-1}^{gross}, 0)$	−0.009* (0.008)	−0.013** (0.004)	−0.064** (< 0.001)	−0.058** (< 0.001)
$\max(R_{t-1}^{gross}, 0)$	0.011** (0.003)	0.008 (0.055)	0.010 (0.226)	0.020** (< 0.001)
$\log(AUM)_{t-1}$	0.025* (0.018)	0.011 (0.322)	−0.033 (0.276)	−0.044** (< 0.001)
EXPERIENCE	−0.003 (0.332)	−0.001 (0.689)	0.014 (0.099)	0.009* (0.035)
TEAM	0.050 (0.369)	0.005 (0.890)	−0.118 (0.153)	−0.100 (0.156)
Constant	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	4,398	9,785	4,517	9,963
Number of managers	58	84	58	84
Adjusted R^2	0.154	0.125	0.620	0.512

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the “Reading-the-Mind-in-the-Eyes”-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the wofo, answered on scales ranging from 1 to 7 each. $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month’s gross return if it was negative; similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month’s gross return if it was positive. $\log(AUM)$ stands for the log of assets under management, EXPERIENCE indicates years in industry, and TEAM is a dichotomous indicator for team-managed funds.

Table S8: Robustness check for performance variables, relative volatility, and tracking error when including time preferences (PATIENCE): The table shows the results of ordinary least squares regressions of funds’ abnormal returns, values added, relative volatility, and tracking error on cognitive skills and economic preferences/attitudes. The abnormal return, R^{abn} , is the total monthly return of the fund over its benchmark. The value added, V , is the product of assets under management, and abnormal return. The relative volatility, RV , stands for the overall riskiness of the fund relative to the riskiness of the benchmark. The tracking error, TE , measures fund’s net return minus the benchmark return. Standard errors are clustered at the manager level. Corresponding p -values are shown in parentheses. ** $p < 0.005$, * $p < 0.05$.

<i>Dependent variable</i>	R^{abn}	V	RV	TE
CRT	0.002 (0.908)	0.040 (0.509)	-0.036* (0.017)	0.003 (0.953)
TOM	-0.003 (0.904)	0.002 (0.965)	0.002 (0.939)	-0.008 (0.811)
COMPETITIVENESS	0.013 (0.660)	0.033 (0.642)	0.019 (0.278)	0.044 (0.329)
RISK TOLERANCE	-0.002 (0.440)	0.008 (0.473)	0.005** (0.005)	0.005 (0.168)
LOSS TOLERANCE	-0.001 (0.811)	0.000 (0.974)	0.001 (0.768)	0.009 (0.111)
AMBIGUITY TOLERANCE	0.002 (0.183)	0.003 (0.402)	-0.001 (0.305)	0.008* (0.029)
PATIENCE	-0.006 (0.257)	-0.017 (0.463)	-0.005 (0.202)	-0.013 (0.058)
$\min(R_{t-1}^{gross}, 0)$			-0.010** (0.002)	-0.070** (< 0.001)
$\max(R_{t-1}^{gross}, 0)$			0.009** (0.001)	0.016* (0.007)
$\log(AUM)_{t-1}$	0.019 (0.129)		0.026** (0.003)	-0.049* (0.008)
EXPERIENCE	0.000 (0.929)	-0.001 (0.910)	0.001 (0.561)	0.007 (0.287)
TEAM	0.071 (0.077)	0.243* (0.030)	0.037 (0.299)	-0.110 (0.075)
Constant	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	12,973	12,973	12,556	12,828
Number of managers	74	74	74	74
Adjusted R^2	0.138	0.080	0.118	0.625

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the “Reading-the-Mind-in-the-Eyes”-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the wofo, answered on scales ranging from 1 to 7 each. $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month’s gross return if it was negative; similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month’s gross return if it was positive. $\log(AUM)$ stands for the log of assets under management, EXPERIENCE indicates years in industry, and TEAM is a dichotomous indicator for team-managed funds.

Table S9: Robustness check for relative volatility, tracking error, and relative semi-volatility when calculating measures based on daily data over half a year (rather than a month as reported in the main text): The table shows the results of ordinary least squares regressions of funds’ relative volatility, tracking error, and semi-volatility on cognitive skills and economic preferences/attitudes. The relative volatility, RV , stands for the overall riskiness of the fund relative to the riskiness of the benchmark. The tracking error, TE , measures fund’s net return minus the benchmark return. Relative semi-volatility, RSV , measures the funds’ down-side risk relative to the downside risk of the benchmark. Standard errors are clustered at the manager level. Corresponding p -values are shown in parentheses. ** $p < 0.005$, * $p < 0.05$.

<i>Dependent variable</i>	<i>RV</i>	<i>RV</i>	<i>TE</i>	<i>TE</i>	<i>RSV</i>	<i>RSV</i>
<i>Sample of funds</i>	<i>All</i>	<i>Single</i>	<i>All</i>	<i>Single</i>	<i>All</i>	<i>Single</i>
CRT	-0.029* (0.020)	-0.041* (0.022)	0.028 (0.532)	0.021 (0.699)	-0.031* (0.030)	-0.045* (0.013)
TOM	-0.009 (0.701)	-0.051* (0.009)	-0.032 (0.397)	-0.023 (0.583)	-0.001 (0.979)	-0.049* (0.012)
COMPETITIVENESS	0.014 (0.381)	0.012 (0.570)	0.066 (0.158)	-0.054 (0.258)	0.021 (0.248)	0.016 (0.437)
RISK TOLERANCE	0.006** (< 0.001)	0.010** (0.002)	0.003 (0.567)	-0.003 (0.489)	0.007** (< 0.001)	0.010** (0.002)
LOSS TOLERANCE	0.002 (0.688)	0.003 (0.499)	0.007 (0.232)	0.009 (0.115)	0.001 (0.848)	0.003 (0.546)
AMBIGUITY TOLERANCE	-0.001 (0.222)	-0.001 (0.215)	0.007* (0.023)	0.009** (0.002)	-0.001 (0.224)	-0.002 (0.206)
$\min(R_{t-1}^{gross}, 0)$	-0.002 (0.198)	-0.001 (0.716)	-0.021** (< 0.000)	-0.019** (0.001)	-0.002 (0.331)	0.000 (0.991)
$\max(R_{t-1}^{gross}, 0)$	0.002 (0.088)	0.002 (0.112)	0.001 (0.704)	0.001 (0.807)	0.003 (0.136)	0.003 (0.087)
$\log(AUM)_{t-1}$	0.015 (0.170)	-0.005 (0.697)	-0.039* (0.024)	-0.064* (0.007)	0.015 (0.156)	-0.004 (0.730)
EXPERIENCE	-0.002 (0.481)	0.000 (0.844)	0.011 (0.052)	0.008 (0.179)	-0.003 (0.470)	0.000 (0.998)
TEAM	0.018 (0.663)		-0.106 (0.119)		0.023 (0.614)	
Constant	yes	yes	yes	yes	yes	yes
Category FE	yes	yes	yes	yes	yes	yes
Time FE	yes	yes	yes	yes	yes	yes
Number of observations	2,170	1,042	2,210	1,081	2,167	1,042
Number of managers	83	51	83	51	83	51
Adjusted R^2	0.155	0.312	0.613	0.674	0.097	0.284

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the “Reading-the-Mind-in-the-Eyes”-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the wofo, answered on scales ranging from 1 to 7 each. $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month’s gross return if it was negative; similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month’s gross return if it was positive. $\log(AUM)$ stands for the log of assets under management, EXPERIENCE indicates years in industry, and TEAM is a dichotomous indicator for team-managed funds.

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

<https://www.uibk.ac.at/eeecon/wopec/>

- 2019-16 **Adam Farago, Martin Holmen, Felix Holzmeister, Michael Kirchler, Michael Razen:** Cognitive Skills and Economic Preferences in the Fund Industry
- 2019-15 **Christopher Kah, Daniel Neururer:** Generiert der stationäre Buchhandel positive Nachfrageeffekte und verhilft dadurch dem Kulturgut Buch bei seiner Verbreitung? - Ein natürliches Experiment
- 2019-14 **Stjepan Srhoj, Michael Lapinski, Janette Walde:** Size matters? Impact evaluation of business development grants on SME performance
- 2019-13 **Andrea M. Leiter, Engelbert Theurl:** DETERMINANTS OF PREPAID SYSTEMS OF HEALTH-CARE FINANCING - A WORLDWIDE COUNTRY-LEVEL PERSPECTIVE
- 2019-12 **Michael Razen, Michael Kirchler, Utz Weitzel:** Domain-Specific Risk-Taking Among Finance Professionals
- 2019-11 **Jonathan Hall, Rudolf Kerschbamer, Daniel Neururer, Eric Skoog:** Uncovering sophisticated discrimination with the help of credence goods markups - evidence from a natural field experiment
- 2019-10 **Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Collective intertemporal decisions and heterogeneity in groups
- 2019-09 **Morten Hedegaard, Rudolf Kerschbamer, Daniel Müller, Jean-Robert Tyran:** Distributional Preferences Explain Individual Behavior Across Games and Time
- 2019-08 **Daniel Müller, Sander Renes:** Fairness Views and Political Preferences - Evidence from a representative sample
- 2019-07 **Florian Lindner, Michael Kirchler, Stephanie Rosenkranz, Utze Weitzel:** Social Status and Risk-Taking in Investment Decisions
- 2019-06 **Christoph Huber, Julia Rose:** Individual attitudes and market dynamics towards imprecision
- 2019-05 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Florian Lindner, Utz Weitzel, Stefan Zeisberger:** What Drives Risk Perception? A Global Survey with Financial Professionals and Lay People
- 2019-04 **David M. McEvoy, Tobias Haller, Esther Blanco:** The Role of Non-Binding Pledges in Social Dilemmas with Mitigation and Adaptation

- 2019-03 **Katharina Momsen, Markus Ohndorf:** When do people exploit moral wiggle room? An experimental analysis in a market setup
- 2019-02 **Rudolf Kerschbamer, Daniel Neururer, Matthias Sutter:** Credence goods markets and the informational value of new media: A natural field experiment
- 2019-01 **Martin Geiger, Eric Mayer, Johann Scharler:** Inequality and the Business Cycle: Evidence from U.S. survey data
- 2018-18 **Matthias Sutter, Jürgen Huber, Michael Kirchler, Matthias Stefan, Markus Walzl:** Where to look for the morals in markets?
- 2018-17 **Rene Schwaiger, Michael Kirchler, Florian Lindner, Utz Weitzel:** Determinants of investor expectations and satisfaction. A study with financial professionals
- 2018-16 **Andreas Groll, Julien Hambuckers, Thomas Kneib, Nikolaus Umlauf:** LASSO-Type Penalization in the Framework of Generalized Additive Models for Location, Scale and Shape
- 2018-15 **Christoph Huber, Jürgen Huber:** Scale matters: Risk perception, return expectations, and investment propensity under different scalings
- 2018-14 **Thorsten Simon, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Lightning prediction using model output statistics
- 2018-13 **Martin Geiger, Johann Scharler:** How do consumers interpret the macroeconomic effects of oil price fluctuations? Evidence from U.S. survey data
- 2018-12 **Martin Geiger, Johann Scharler:** How do people interpret macroeconomic shocks? Evidence from U.S. survey data
- 2018-11 **Sebastian J. Dietz, Philipp Kneringer, Georg J. Mayr, Achim Zeileis:** Low visibility forecasts for different flight planning horizons using tree-based boosting models
- 2018-10 **Michael Pfaffermayr:** Trade creation and trade diversion of regional trade agreements revisited: A constrained panel pseudo-maximum likelihood approach
- 2018-09 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model
- 2018-08 **Lisa Schlosser, Torsten Hothorn, Reto Stauffer, Achim Zeileis:** Distributional regression forests for probabilistic precipitation forecasting in complex terrain
- 2018-07 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Delegated decision making and social competition in the finance industry
- 2018-06 **Manuel Gebetsberger, Reto Stauffer, Georg J. Mayr, Achim Zeileis:** Skewed logistic distribution for statistical temperature post-processing in mountainous areas

- 2018-05 **Reto Stauffer, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach
- 2018-04 **Utz Weitzel, Christoph Huber, Florian Lindner, Jürgen Huber, Julia Rose, Michael Kirchler:** Bubbles and financial professionals
- 2018-03 **Carolin Strobl, Julia Kopf, Raphael Hartmann, Achim Zeileis:** Anchor point selection: An approach for anchoring without anchor items
- 2018-02 **Michael Greinecker, Christopher Kah:** Pairwise stable matching in large economies
- 2018-01 **Max Breitenlechner, Johann Scharler:** How does monetary policy influence bank lending? Evidence from the market for banks' wholesale funding

University of Innsbruck

Working Papers in Economics and Statistics

2019-16

Adam Farago, Martin Holmen, Felix Holzmeister, Michael Kirchler, Michael Razen

Cognitive Skills and Economic Preferences in the Fund Industry

Abstract

By running a battery of incentivized and non-incentivized experiments with fund managers from four countries in the European Union, we investigate the impact of fund manager' cognitive skills and economic preferences on the dynamics of the mutual funds they manage. First, we find that fund managers' risk tolerance positively correlates with fund risk when accounting for fund benchmark, fund category, and other controls. Second, we show that fund managers' ambiguity tolerance positively correlates with the funds' tracking error from the benchmark. Finally, we report that cognitive skills do not explain fund performance in terms of excess returns. However, we do find that fund managers with high cognitive reflection abilities generate these returns at lower risk.

ISSN 1993-4378 (Print)

ISSN 1993-6885 (Online)