I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11327

Confidence in Knowledge or Confidence in the Ability to Learn: An Experiment on the Causal Effects of Beliefs on Motivation

Mira Fischer
Dirk Sliwka

FEBRUARY 2018

DISCUSSION PAPER SERIES

IZA DP No. 11327

Confidence in Knowledge or Confidence in the Ability to Learn: An Experiment on the Causal Effects of Beliefs on Motivation

Mira Fischer
University of Cologne

Dirk Sliwka
University of Cologne and IZA

FEBRUARY 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Confidence in Knowledge or Confidence in the Ability to Learn: An Experiment on the Causal Effects of Beliefs on Motivation*

Previous research has shown that feedback about past performance has ambiguous effects on subsequent performance. We argue that feedback affects beliefs in different dimensions – namely beliefs about the *level of human capital* and beliefs about the *ability to learn* – and this may explain some of the ambiguous effects. We experimentally study the causal effects of an exogenously administered change in beliefs in both of these dimensions on the motivation to learn. We find that confidence in the ability to learn raises incentives, while confidence in the level of human capital lowers incentives for individuals with high levels of human capital.

JEL Classification: C91, D83, I21, J24

Keywords: economic experiments, confidence, human capital investment, motivation

Corresponding author:
Dirk Sliwka
University of Cologne
Albertus-Magnus-Platz
50923 Cologne
Germany
E-mail: dirk.sliwka@uni-koeln.de

* Financial support by the Deutsche Forschungsgemeinschaft (DFG) through Research Unit FOR 1371 and by the University of Cologne through the Hans Kelsen Research Prize is gratefully acknowledged.

"Take two people in school: One of them is a know-it-all and the other is a learn-it-all. Even if the know-it-all starts with more innate capability, the learn-it-all will ultimately outperform him."

Satya Nadella, CEO of Microsoft, in June 2016

1 Introduction

Motivational beliefs are held to be a strong determinant of important life outcomes such as educational attainment and professional development. However, there seems to be disagreement in the public realm on which beliefs about ourselves are beneficial for us. Folk wisdom tells us that holding a very favorable opinion of our abilities may often breed failure as it tempts us to rest on our laurels and lowers our motivation to work hard towards our goals and the economics literature, too, mostly emphasizes the negative effects of too much confidence. However, many popular self-help books claim that increasing our self-confidence makes us more likely to be successful in life.[1] In educational settings, optimistic beliefs about ourselves are widely thought to foster skill development and a quick search on the internet will turn up many school-related websites and workshop offers claiming that fostering children's confidence will improve their motivation to learn. However, there seems to be disagreement about whether praise for performance, effort, or progress is best to raise confidence and motivation to learn.

A straightforward conjecture is that some of the disagreement in the popular discourse about the relationship between feedback, confidence, and performance is caused by the tendency to subsume different types of beliefs under the notion of "confidence". Different types of feedback may influence beliefs about different dimensions of a person's skills and abilities and conditional on circumstances a shift in a belief about a given skill dimension may or may not raise motivation to exert effort.[2]

The key purpose of this paper is to distinguish two dimensions of confidence – *confidence in one's level of prior knowledge* and *confidence in one's learning ability* – and to study causal effects of changes in these dimensions of a person's confidence on investments in human capital. Reinforcement of confidence in these two dimensions likely has very different effects, as the

[1] The claim "confidence breeds success" produces 329 hits on Google Books and a search on Amazon.com for "confidence" in the sub-category "Books - Self-Help - Success" produces 783 hits.

[2] Indeed, the literature in psychology indicates that there is mixed evidence on the association between different types of feedback and performance (Kluger and DeNisi, 1998; Hattie and Timperley, 2007).

first dimension is related to one's ex-ante probability of passing a test while the second one is related to how much one's passing probability increases when exerting learning efforts. We first illustrate these belief dimensions in a simple formal model and then study the effects of exogenous variation in both dimensions in a lab experiment.

The motivational role of confidence has attracted substantial interest from different fields in economics in recent years. Bénabou and Tirole (2002, 2003), for instance, have studied formal models in which agents are uncertain about the marginal returns to their effort. These models yield a precise notion of confidence as an agent's belief in her own marginal product of effort. A higher confidence then naturally induces an agent to work harder on a task.[3] The recent literature on the economics of education has studied specific personality traits that predict important life outcomes (Heckman et al., 2006; Cebi, 2007; Heineck and Anger, 2010; Heckman and Kautz, 2012). Internal locus of control and self-esteem, psychological constructs intended to capture a person's beliefs about the ability to affect outcomes, feature prominently among these traits. There is also empirical evidence that socially disadvantaged children (Filippin and Paccagnella, 2012), and girls (Reuben et al., 2013) are less confident about their academic ability and that this has negative effects on their educational decisions and expected earnings.

In our experiment students have to decide how intensively they want to prepare for a test. They pass the test and earn a reward if their performance reaches a certain threshold. Based on the analysis of a simple formal model we hypothesize that a higher *confidence in the level of prior knowledge* causes students with low levels of knowledge to invest more. This is because it subjectively moves them closer to the passing threshold and raises the probability that an additional remembered item is pivotal to passing the test. For students with high levels of prior knowledge we expect the opposite, i.e. that raising their confidence in knowledge even further will lower their effort to prepare for the test because it subjectively moves them further away from the passing threshold such that learning becomes less relevant for whether someone passes or fails the test. For the other dimension – *confidence in learning ability* – we expect that raising this dimension of confidence will have a monotonic effect and cause students to invest more effort in learning because the perceived marginal cost of effort to generate "knowledge" decreases.

To study the causal effects of the two dimensions of confidence, we exogenously vary feedback scores subjects receive about their performance in two prior tests. One of these tests measures their prior knowledge, the other test measures the ability to memorize information. After com-

[3] See, for instance, Koch et al. (2015) for an overview on these and related models from the perspective of the economics of education.

pleting these two tests, each subject privately receives a feedback score for each of the tests. Subjects know that each feedback score is the sum of their true score in the respective test and a random noise term. We then elicit subjects' confidence by asking them to estimate their own rank in the first two tests. Subjects can then buy pieces of information and memorize these to prepare for a final test in which they earn a fixed amount of money if their performance exceeds a specific threshold. The random component in the feedback scores thus generates exogenous variation in the agents' confidence in the two dimensions, which we use as instrumental variables to estimate causal effects of confidence on investment decisions and test outcomes.

We find that a higher confidence in learning raises learning investments irrespective of the prior level of knowledge. Confidence in knowledge, however, has a negative effect on investments of individuals with above average prior knowledge and a positive effect on investments of individuals with below average prior knowledge. With respect to test outcomes, we find that raising the confidence in learning of individuals with below average prior knowledge improves their rank in the final test and their probability of passing it, however, we do not find a beneficial effect for individuals who already had above average prior knowledge. Mirroring the effects of confidence in knowledge on effort, we find that raising confidence in knowledge of individuals with above average prior knowledge decreases their outcomes in the final test whereas it has the opposite effect on individuals with below average prior knowledge.

This paper makes two contributions. First, it shows theoretically and experimentally that in situations where choices involve effort, confidence should be viewed as a multidimensional concept (even if the effort choice is unidimensional) and that general statements about the motivational effects of confidence are misleading. In order to explain the effects of confidence on motivation to exert effort, and on learning in particular, we have to understand which roles effort and ability play in achieving a goal. An important implication of this is also that interventions aimed at raising confidence should be carefully designed and evaluated because they might affect several beliefs that interact in different ways with motivation to exert effort. Second, we develop a deception-free experimental approach to study the *causal* effect of beliefs on effort by generating exogenous variation in two dimensions of confidence. For this reason, we can rule out that, for instance, unobserved psychological dispositions that may be correlated with confidence drive the association between confidence, motivation to exert effort, and performance. By studying the effects of confidence on learning decisions and test outcomes, our study links the literature on experiments in education to the literature on motivational beliefs and socio-emotional skills.

The remainder of the paper is structured as follows. Section 2 summarizes the related literature on the determinants of effort provision in educational and similar settings. Section 3 presents a model and derives best responses and hypotheses from it. Section 4 presents the experimental design. Section 5 presents the results and Section 6 concludes.

2 Related Literature

Our research is closely related to the game theoretical and behavioral economic literature on confidence and incentives. As stated above, "confidence in learning ability" in our setting is equivalent to Benabou and Tirole's (2002, 2003) notion of confidence as an agent's (rational) belief in her own marginal product of effort. We study the interplay between this type of confidence and confidence in prior knowledge as well as the impact of both on investment incentives.[4]

The effects of beliefs in and feedback about ability have been explored in several theoretical papers. The role of feedback in tournament settings has, for instance, been explored by Aoyagi (2010) and Gershkov and Perry (2009). Most closely related to our study is the analysis by Ederer (2010) who studies the effect of interim feedback (about interim outcomes) on effort and shows that when effort and ability are complements feedback should induce competing effects as it informs agents about their relative standing (which reduces incentives) as well as their ability (which may increase incentives). In a principal-agent setting, Santos-Pinto (2008) shows that a worker's overestimation of his ability is beneficial for the principal when ability and effort are complements but not when they are substitutes. Our experiment provides causal empirical evidence for the relevance of disentangling different ability beliefs.

In the context of job search on the labor market, contributions by Caliendo et al. (2015) and Spinnewijn (2015) have studied the role of different dimensions of confidence on search efforts. Most closely related to our model is the analysis of Spinnewijn (2015), who studies how biased beliefs in two dimensions influence job search: "baseline beliefs" – the beliefs about the baseline job finding probability for given search efforts, and "control beliefs" – the beliefs about the increase in the job finding probability when searching more intensively. We study the effect of baseline belief (concerning prior knowledge) and control belief (concerning ability to learn) on learning effort and provide causal evidence on their impact in an educational setting.

A number of empirical and experimental papers have studied the effect of feedback about

[4] Compte and Postlewaite (2004) depart even further from a neoclassical framework by assuming that confidence, influenced by an agent's past successes and failures, raises the (factual) probability of success of an agent.

(relative) performance on educational outcomes. Tran and Zeckhauser (2012) find that students perform significantly better in a final English test when they are told their rankings on practice tests than students in the control group who only receive private feedback about their test score. Bandiera et al. (2015) exploit rule differences between university departments concerning the provision of feedback to students and find that students who receive their individual exam grade prior to writing a long essay do better in it than students who do not. Azmat and Iriberri (2010), in a natural field experiment set in a high school, find that students who repeatedly receive information about the average grade of their class in addition to information about their own grade, receive 5 percent better grades. In Azmat et al. (2015), however, a random sample of college students who receive information about their position in the distribution of grades repeatedly over a period of three years are found to do worse during the first six months. As the authors argue, students in their sample were initially underconfident. Thus learning that they were doing better than expected had a negative impact on performance. In line with this argument, Kuhnen and Tymula (2012), who study effort reactions to rank feedback in the lab, find that individuals who ranked better than expected decrease output, whereas those who ranked worse than expected increase output. In contrast to these studies, we do not vary feedback on the relative rank in the relevant test but go one step back and manipulate the beliefs a person holds about her knowledge and ability to learn in order to shed light on the behavioral mechanisms by which feedback affects behavior.

Finally, although incentive compatible measurement of beliefs is common in economic laboratory studies, there are very few studies which generate *exogenous* variation in beliefs in order to study the causal effect of beliefs on actions. Mobius et al. (2011) repeatedly give noisy feedback about whether one performed in the better or the worse half of participants in an IQ test. The authors use the random variation in the feedback to estimate the causal effect of confidence in own ability on the aversion to receiving information about ability and find that a lower confidence induces a stronger aversion to receiving information about one's own ability. Schwardmann and Van der Weele (2016) investigate the hypothesis that overconfidence serves to more effectively persuade others and also manipulate subjects' confidence in their own intelligence using noisy feedback. Costa-Gomes et al. (2014) study the causal effect of beliefs in a trust game by inducing a zero-mean random shift that exogenously increases or reduces the trustee's level of re-payment. Then the authors use the random shift as instrumental variable to estimate the causal effect of beliefs about the trustee's transfer share on the trustor's choice.

Our study is the first that uses noisy feedback to manipulate two different belief dimensions in order to study the causal effect of ability beliefs on learning investments and test outcomes.

3 An Illustrative Model

Consider the following simple illustrative model which can be interpreted as an analysis of a reaction function in a standard Lazear and Rosen (1981) tournament in which we allow the agent's beliefs to vary with respect to (i) the costs of effort (ability a) and (ii) a potential handicap/or lead (prior knowledge k). In contrast to the standard tournament literature we do not analyze the equilibrium behavior of a small set of players but follow Casas-Arce and Martínez-Jerez (2009) in studying a "population tournament" where the threshold necessary to win the prize is deterministic. The model's purpose is to illustrate how changes in these two forms of "confidence" should affect the efforts exerted to win the prize.

A risk neutral agent can invest effort to raise her human capital. Human capital is measured by "pieces of knowledge". An agent's posterior knowledge is the sum of her prior knowledge k and knowledge acquired through learning Δ. Knowledge acquisition is costly and the agent's cost function is

$$c\left(\Delta, a\right)$$

where a measures the agent's ability to acquire further knowledge. We assume that $\frac{\partial c}{\partial \Delta}, \frac{\partial^2 c}{\partial \Delta^2} > 0$ and $\frac{\partial c}{\partial \Delta \partial a} < 0$ such that the marginal costs of knowledge acquisition are smaller for more able agents. The agent is uncertain about both, her prior knowledge k and the ability to acquire further knowledge a. She knows that both are distributed according to the cumulative distribution functions $F_a\left(a\right)$ and $F_k\left(k\right)$. The agent receives informative signals $s = \left(s_a, s_k\right)$ such that $\frac{\partial E[a|s_a, s_k]}{\partial s_a} > 0$ and $\frac{\partial E[k|s_a, s_k]}{\partial s_k} > 0$. Note that we can decompose

$$\begin{aligned} a &= E\left[a\left|s_a, s_k\right.\right] + \varepsilon_{as} \\ k &= E\left[k\left|s_a, s_k\right.\right] + \varepsilon_{ks} \end{aligned}$$

where ε_{as} and ε_{ks} are uncorrelated with the signals $\left(s_a, s_k\right)$ and have mean zero (by the law of iterated expectations).[5] Assume that ε_{as} and ε_{ks} have unimodal densities with $g'_{\varepsilon_{as}}\left(0\right) =$

[5] To see, for instance, that $Cov\left[s_a, \varepsilon_{as}\right] = Cov\left[s_a, a - E\left[a\left|s_a, s_k\right.\right]\right] = 0$ note that by the law of iterated expectations $E\left[s_a\left(a - E\left[a\left|s_a, s_k\right.\right]\right)\right] = E\left[E\left[s_a\left(a - E\left[a\left|s_a, s_k\right.\right]\right)\left|s_\gamma, s_k\right.\right]\right] = E\left[s_a E\left[\left(a - E\left[a\left|s_a, s_k\right.\right]\right)\left|s_a, s_k\right.\right]\right] = 0$.

$g'_{\varepsilon_{ks}}(0) = 0$. Denote the conditional expectations as

$$\begin{aligned} \hat{k} &= E[k|s_a, s_k] \\ \hat{a} &= E[a|s_a, s_k] \end{aligned}$$

such that $\hat{k}$ and $\hat{a}$ describe the agent's own mean belief in her knowledge and costs of knowledge acquisition respectively. The decomposition allows us to do comparative statics with respect to $\hat{k}$ and $\hat{a}$, which capture an agent's confidence in the two dimensions.

The agent attains a certain educational outcome, such as passing an admission test to an education program, or being awarded an academic title, if $k + \Delta$ exceeds a threshold value τ.[6] In this case she will receive a reward B. The agent's objective function can thus be denoted as

$$\max_{\Delta} \Pr\left(\hat{k} + \varepsilon_{ks} + \Delta > \tau\right) B - E\left[c(\Delta, a)|s_a, s_k\right].$$

In order to guarantee that this optimization problem has a unique solution we assume that

$$\max_{\varepsilon}\left(-g'_{\varepsilon_{ks}}(\varepsilon)\right) B < \min_{\Delta,a} E\left[\frac{\partial^2 c(\Delta, a)}{\partial \Delta^2}\right] \tag{1}$$

which will, for instance, hold if $\frac{\partial^2 c(\Delta,a)}{\partial \Delta^2}$ is bounded from below by a constant and the signal s_k is not too precise.[7]

The first derivative of the objective function is

$$g_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) B - E\left[\frac{\partial c(\Delta, \hat{a} + \varepsilon_a)}{\partial \Delta}\right]$$

and by condition (1) the objective function is strictly concave. We can now show:

Proposition 1 *Knowledge acquired through learning $\Delta\left(\hat{a}, \hat{k}\right)$ is strictly increasing in the agent's confidence in her ability to acquire knowledge $\hat{a}$. It is strictly increasing in the agent's confidence in prior knowledge $\hat{k}$ if and only if $\hat{k}$ is smaller than a cut-off value and otherwise strictly decreasing.*

[6] Note that here we treat τ as an exogenous constant. If we consider a tournament setting τ will be determined in equilibrium by the choices of the other agents. In a tournament between a continuum of agents where a fixed fraction can win a prize the equilibrium threshold will indeed be deterministic (see, for instance Casas-Arce and Martínez-Jerez (2009)).

[7] This condition will guarantee that the objective function is strictly concave. Intuitively, if there is sufficient uncertainty on k then ε_{ks} will have a large variance. If, for instance, ε_{ks} is normally distributed a large enough variance will guarantee that the slope of the density function will not be too steep.

Proof:

By implicit differentiation we obtain

$$\frac{\partial \Delta\left(\hat{a}, \hat{k}\right)}{\partial a} = -\frac{-E\left[\frac{\partial c(\Delta, \hat{a}+\varepsilon_a)}{\partial \Delta \partial a}\right]}{-g'_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) B - E\left[\frac{\partial^2 c(\Delta, \hat{a}+\varepsilon_a)}{\partial \Delta^2}\right]} > 0$$

as the denominator is negative by condition (1). And

$$\frac{\partial \Delta\left(\hat{a}, \hat{k}\right)}{\partial \hat{k}} = -\frac{-g'_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) B}{-g'_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) B - E\left[\frac{\partial^2 c(\Delta, \hat{a}+\varepsilon_a)}{\partial \Delta^2}\right]} \tag{2}$$

such that

$$\frac{\partial \Delta\left(\hat{a}, \hat{k}\right)}{\partial \hat{k}} > 0 \Leftrightarrow g'_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) < 0$$

which, as $g_{\varepsilon_{ks}}(\varepsilon)$ has a unique mode at 0, is equivalent to

$$\tau > \hat{k} + \Delta\left(\hat{a}, \hat{k}\right).$$

The right hand side is strictly increasing k as $\frac{\partial \Delta(\hat{a}, \hat{k})}{\partial k} > -1$. To see the latter, note that

$$\frac{\partial \Delta\left(\hat{a}, \hat{k}\right)}{\partial \hat{k}} = -\frac{-g'_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) B}{-g'_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) B - E\left[\frac{\partial^2 c(\Delta, \hat{a}+\varepsilon_a)}{\partial \Delta^2}\right]} > -1 \Leftrightarrow$$

$$g'_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) B < g'_{\varepsilon_{ks}}\left(\tau - \hat{k} - \Delta\right) B + E\left[\frac{\partial^2 c\left(\Delta, \hat{a} + \varepsilon_a\right)}{\partial \Delta^2}\right]$$

which always holds. Hence, condition (2) holds for sufficiently small k and will not hold above a threshold level.[8] ■

To illustrate the result, consider the following parametric example. Assume that the agent's cost function is $c(\Delta, a) = \frac{c-a}{2}\Delta^2$ and that the agent believes that k is normally distributed with mean $\hat{k}$ and variance $V[\varepsilon_{ks}] = \sigma^2_{\varepsilon_k}$. As the cost function is linear in a, expected costs are equal to $\frac{c-\hat{a}}{2}\Delta^2$. The agent's objective function is thus

$$\max_{\Delta} \Pr\left(\varepsilon_{ks} > \tau - \Delta - \hat{k}\right) B - \frac{c - \hat{a}}{2}\Delta^2.$$

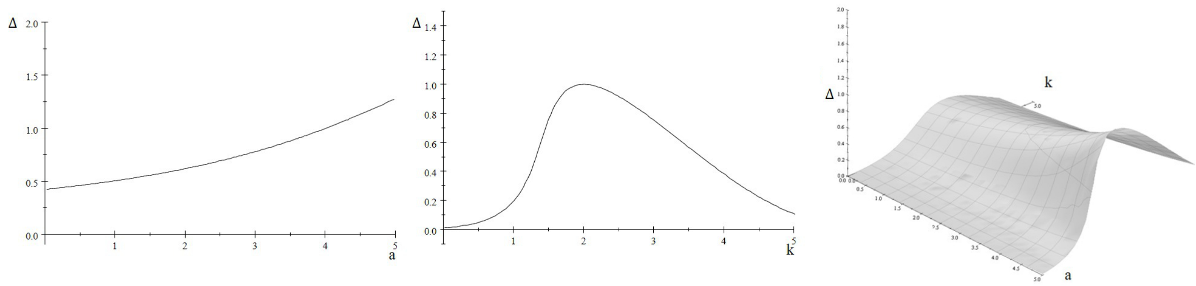
[8]Note that this threshold will be strictly positive if $\tau > k + \Delta\left(\hat{a}, \hat{k}\right)$ for $k = 0$. A sufficient condition for this is that the objective function is downward sloping in Δ at $\Delta = \tau$ for $k = 0$, which is the case when $g_{\varepsilon_{ks}}(0) B < E\left[\frac{\partial c(\Delta, \hat{a}+\varepsilon_a)}{\partial \Delta}\right]$. This will hold if the signal on k is not too precise.

The first derivative of the objective function[9] becomes

$$\frac{1}{\sigma_{\varepsilon_k}}\phi\left(\frac{\tau-\Delta-\hat{k}}{\sigma_{\varepsilon_k}}\right)B-(c-\hat{a})\,\Delta=0,$$

where $\phi(\varepsilon)$ is the pdf of a standard normal distribution. While this equation has no closed form solution we can use this expression to plot Δ as an implicit function of a and Δ for specific examples.[10]

Figure 1: Learning investments as a function of perceived ability and knowledge



Hence, *a higher confidence in the ability to learn always leads to higher learning investments* as it lowers the perceived marginal costs of learning efforts. This is essentially the motivational effect of self-confidence stressed, for instance, by Benabou and Tirole (2002). However, *confidence in prior knowledge* has a positive effect only for agents with low prior knowledge but reduces the incentives to learn for those with higher prior knowledge. The intuition is the following: If an agent has rather low confidence in her initial knowledge she thinks that the likelihood of achieving the educational outcome is small. In turn, the expected marginal gains from learning are small. Raising the confidence in knowledge raises the perceived likelihood to jump the threshold and consequently increases the marginal returns to learning efforts. If, however, the agent believes that she has a very high level of prior knowledge, her perceived likelihood of attaining the outcome even at lower learning investments increases. In turn, the incentive to invest in acquiring further knowledge decreases.

Based on this illustrative model, we designed an experiment that enables us to clearly disentangle confidence in prior knowledge and confidence in the ability to learn and allows us to

[9] Condition (1) that guarantees an internal solution here becomes $\frac{1}{\sigma^2_{\varepsilon_k}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}}B < c-\hat{a}$, i.e. the objective function will be strictly concave if $\hat{a}$ is not too large.

[10] The plots use values $B=10$, $\sigma^2_{\varepsilon_k}=1$,
$\tau=3$, $c=8$ and the condition guaranteeing a strictly concave objective function requires that $\hat{a}<8-\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}}10=5.5803$.

measure the causal effect of confidence in both dimensions.

4 Experimental Design

We have to keep in mind that confidence is inherently an endogenous variable as it will always be affected by unobserved experiences, abilities, and other traits of the respective subjects, which could also affect the outcome variables through different unobserved behavioral channels. Hence, merely detecting a correlation between confidence and behavior does not allow to infer causality. In order to avoid this problem, we have developed an experimental design in which we generate *instrumental variables*, that is variables that are (i) cleanly exogenous but (ii) directly affect confidence. We then use these variables to investigate the causal effects of confidence on behavior. In the following will explain in detail how we implemented this idea.

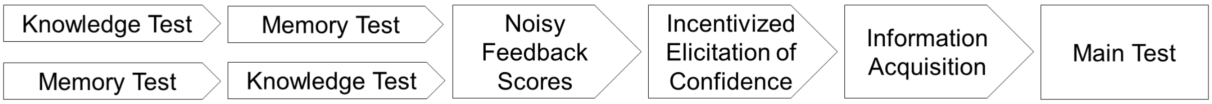
We invited university students to the Cologne Laboratory for Economic Research.[11] Upon arrival, registered participants were randomly assigned a computer. Before the experiment started, students were informed that they were prohibited to talk to each other, to use electronic devices or pen and paper during the experiment and that anyone who violated this rule would be excluded from the experiment. We monitored compliance with the rule during the entire session. Participants were informed that they would receive the regular show-up fee of 2.50 euros and that they could earn additional money during the experiment.[12]

The timeline of the experiment is illustrated in figure 2 and can be summarized as follows: Before the main intervention, subjects take part in a memory and a knowledge test. Then they learn a feedback score about their performance in each test and these feedback scores are the sum of the respective test outcomes and random noise terms. Hence, this stage constitutes our treatment variation: The noise terms exogenously vary information that should affect subjects' confidence in the two dimensions. In a next step we elicit subjects' beliefs about their relative standing in both domains, which was incentivized by paying them for accuracy of beliefs. These are the main belief variables we use in our analysis as measures of confidence in the two dimensions. Then subjects can undertake a costly investment in further knowledge to prepare for a final test in which they can earn a substantial amount of money when passing a threshold. The learning investment as well as the test results will constitute our outcome variables.

[11] The laboratory uses the recruitment software ORSEE (Greiner, 2004) for managing the subject pool. The experiment was programmed using z-Tree (Fischbacher, 2007)

[12] A detailed description of the experiment's timeline, tests, feedback, and belief elicitation can be found in appendices E and F.

Figure 2: Timeline of the Experimental Procedure



4.1 Stages of the Experiment

Measurement of Prior Knowledge and Learning Ability: After the introduction, participants saw a description of the test they were about to take first, which was either a *"knowledge test"* or a *"memory test"*. The order of the tests was randomized within each session to eliminate possible order effects. In the *knowledge test* subjects had to rank 60 cities according to their numbers of inhabitants within triples of cities, i.e. they had to state which city is the largest and which one is the smallest among three cities and would earn a piece rate of 0.10 euros for each correct set. In the *memory test* subjects first saw a list of 36 cities with a (fictitious) city code belonging to each city. This list was displayed on the screens for 15 minutes and subjects were not allowed to take notes. After this they had to rank cities within triples according to these city codes and would earn 0.20 euros for each correct set. Hence, the knowledge test measured subjects' prior knowledge and the memory test measured their capacity to memorize information. The memory test closely resembles tests used by psychologists to test working memory capacity (Wilhelm et al., 2013) and was designed such that it covers the same domain (numbers attached to city names) as the knowledge test and in order to make one's performance in it seem as relevant as possible with respect to one's later learning decision for a test in this domain.[13]

Both tests were incentivized with a piece rate. Participants took the two tests one after another and after each test were asked how many triples they believed to have solved correctly, immediately afterwards they were also asked how many triples they believed other participants on average solved correctly. In both cases answers were not incentivized and participants were informed that their answer did not have any effect on the further course of the experiment. A detailed overview of the tests and stages of the experiment can be found in Appendices E

[13] Working memory capacity is a strong predictor of ability to acquire knowledge and new skills, independently of IQ (Alloway and Alloway, 2010). See Ackerman et al. (2005) for an overview.

and F.[14] Then participants were informed there will be a "Test 3 (main test)", and that, unlike in the first two tests, they would earn 10 euros if they performed better than half of participants in the session who did the tests in the same order as them. They were also informed that they could prepare for this third test.

Feedback stage: Participants were informed that before preparing for the third test, they would receive feedback about their outcomes in the first two tests in the form of a *"knowledge score"* and a *"memory score"*. As explained to the participants, each score was the sum of a participant's number of correct sets in the respective test and a noise term uniformly and independently distributed between -2 and +2 such that each of the values (-2, -1, 0, 1, 2) is drawn with a probability of 20 percent and added to the true score.[15] The *randomly distributed noise term* thus creates exogenous variation in feedback about knowledge and learning ability while avoiding any form of deception. Then the personal feedback scores and average feedback scores of participants in past sessions were displayed on the same screen.[16] As already noted above, the exogenous variation in the personal feedback scores allows for the estimation of causal effects of the agents' confidence on behavior, a central contribution of our study, and thus an important design feature of our experiment.

Measurement of confidence: Participants were asked to estimate their rank in the knowledge and in the memory test relative to those participants in the room who worked on the two tests in the same order as them. They were informed that they could earn one euro, respectively, for estimating their rank in each test correctly.[17] Our design thus allows us to measure both the perceived level of ability (which is the focus of many economic studies of situations where a choice does not entail a decision about effort), and the perceived effectiveness of effort to raise the level of ability (the focus mainly of psychological studies employing non-

[14]We measured beliefs twice. Once before giving feedback (unincentivized) and once afterwards (incentivized - see details below). Note that the beliefs elicited after the feedback intervention are crucial for our design, as they serve as a measure of confidence that can be affected by the treatment intervention (i.e. the noisy feedback). The unincentivized belief elicitation before giving feedback only has a diagnostic function and allows us to observe the magnitude of subjects' uncertainty about their performance in the two tests. As described further below, uncertainty about own ability and knowledge is generally high prior to learning the test score – which is a precondition for our feedback manipulation to work.

[15]For a similar approach compare, for instance, Grossman and Owens (2012) who study agent's reactions to noisy feedback about their own performance. Note that the incentives in rank order tournaments are not affected by random noise (For a summary of the literature see Dechenaux et al. (2015).)

[16]We always displayed the same average results from a pilot study to keep the frame of reference of the personal feedback constant between the experimental sessions. Participants in the pilot study were recruited from the same subject pool as participants in the experiment and results were very similar.

[17]This method is easy to explain and elicits the mode of an agent's subjective beliefs in an incentive compatible manner and is robust to risk aversion. To see that, note that an agent who has to state an estimate r, the value of a random variable x, and receives 1 euro when reporting correctly should report $argmax_r Pr\left(r=x\right)u(1) + (1-Pr\left(r=x\right))u(0)$, which is equal to the mode of the distribution. Since the range of beliefs in our context is small due to a limited number of ranks, the chances of having an exact estimate are reasonable.

incentivized questionnaires to measure self-efficacy and locus of control (Eccles and Wigfield, 2002)) in an incentive compatible manner.

Investment stage: After participants learned their knowledge score and their memory score they were shown a screen explaining the main *"combined knowledge and memory test"* in detail. Participants were informed that this test was based on the same field of knowledge and had the same length and structure as the initial knowledge test, i.e. they would have to rank sets of three cities according to the size of their populations. This time, however, they would earn a prize of 10 euros when doing better in this test than half of participants in the session who did the first two tests in the same order as them. Furthermore, they were told that they could prepare for it by acquiring information relevant to pass the test. To be specific, subjects had a budget of 3 euros to buy information about cities' numbers of inhabitants in packages of 10 cities for 0.5 euros per package. They could buy a maximum of 6 packages, together covering all the cities in the test. The decision on how many packages to buy was a one-shot decision, i.e. subjects had to state in advance how many packages they wanted to acquire [18] They knew that all cities they could "buy" were part of the later test and each package – when fully memorized – would allow to completely answer at least 3 assignments (triples) in the later test. The acquired packages were then displayed in a 15 minutes learning phase before the final test. In this phase subjects also had the possibility to click on a button in order to look at cartoons displayed on the screen (and subjects knew this before they acquired information).[19] Hence, subjects faced two kinds of costs of learning, direct (and measurable) monetary costs for buying information and (unobservable) mental costs of memorizing the information displayed on the screen.

Final test: Finally, participants took the third *combined knowledge and memory test* in which they had to rank sets of three cities according to the size of their populations. The test is not a pure knowledge test as it includes many smaller cities where a prior pilot has shown that even very knowledgeable subjects may not be able to rank all tuples perfectly without further acquired knowledge from the investment stage. The key idea of the third test is that both, prior knowledge of geography and knowledge acquired during the experiment matter for success. Subjects earned 10 euros if they performed better than the average of participants in the session who did the tests in the same order as them.

After the test, participants filled in a questionnaire. In the very end they were informed

[18]The part of the budget that was not spent, was added to the payoff in the end of the experiment and subjects were aware of this.

[19]This provided them with a task when they finished memorizing or wanted to take a break and induced some opportunity costs of effort.

about how much money they had earned (and how they had performed) in each stage of the experiment.

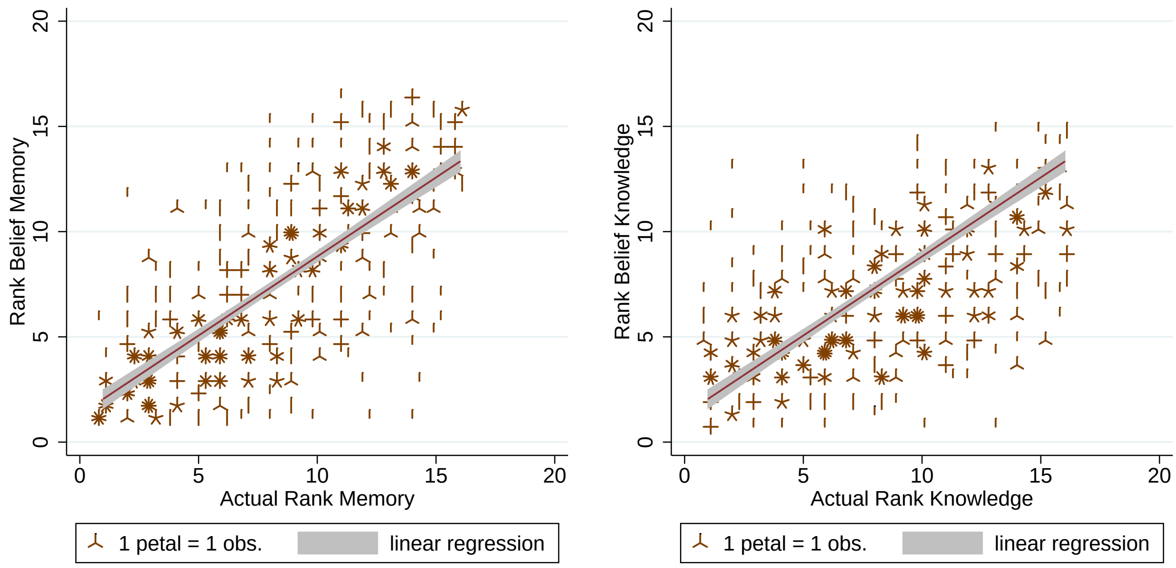
5 Experimental Results

Our main interest is in the size of the learning investment that participants make to prepare for the final test and how this investment is causally affected by confidence in gains and confidence in levels, i.e. beliefs about learning ability and prior knowledge. The key hypotheses are: (i) confidence in the ability to learn should raise learning efforts irrespective of the prior level of knowledge and (ii) confidence in knowledge should increase the incentives to learn for subjects with low prior knowledge and decrease incentives for subjects with high prior knowledge. We measure confidence as agents' beliefs about their relative rank in the memory and knowledge tests elicited after they have learned the respective feedback scores. We ran 16 experimental sessions in May and June and 8 sessions in October 2015. In total 645 people participated in them.[20] The average total payoff was 11.29 euros (including a 2.50 euros show-up fee), the standard deviation of payoffs was 5.01 euros. Subjects on average earned 1.03 euros in the memory test, 0.89 euros in the knowledge test and 5.00 euros in the final test. Sessions lasted approximately one hour and 10 minutes. 63 percent of participants were female. All participants were university students. The mean semester they were in was 6.5.

[20]Instrumental variable regressions allow us to estimate the causal effect of beliefs on behavior but come along with a substantial loss in statistical power (see, for instance, Cameron and Trivedi (2005, section 4.9.3)) the extent of which is hard to gauge in advance without prior knowledge about the variance in the respective test scores and the outcome variable. For this reason we decided to run additional sessions in October 2015 to collect more observations. We can use 615 observations in our estimates as we have some missing data due to cases in which subjects did not submit their answers.

5.1 Descriptive Analysis

Figure 3: Actual Ranks Versus Rank Beliefs

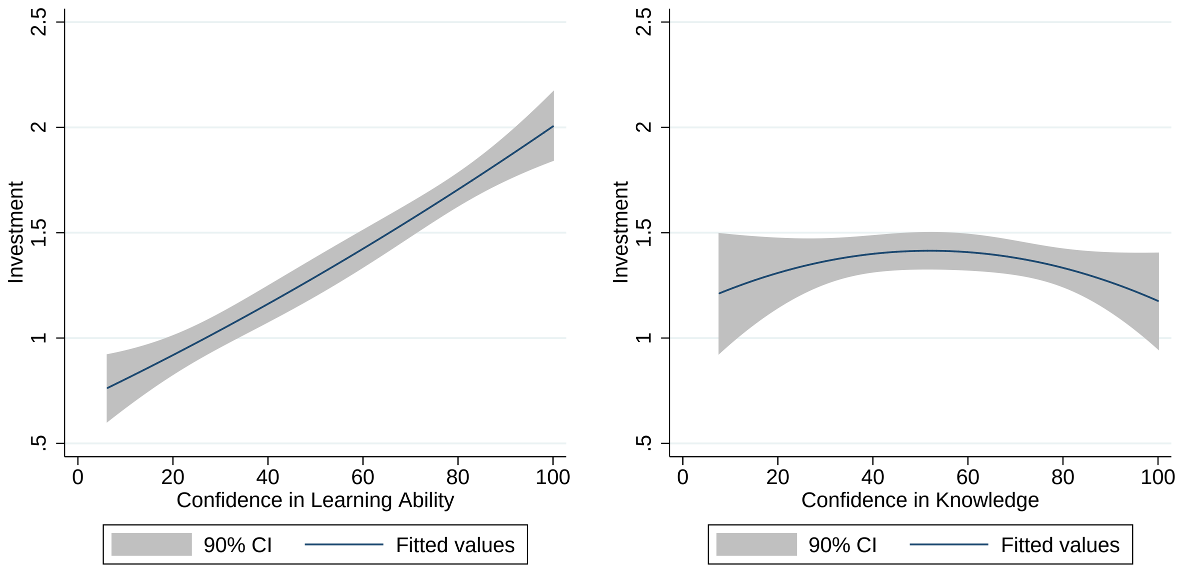


Note: This figure shows ordinal ranks versus rank beliefs with respect to the memory and the knowledge test elicited after giving feedback (1 is best).

We begin by descriptively studying the relationship between rank beliefs *elicited after the feedback intervention* and actual ranks as well as the correlation between these beliefs and investment behavior. The sunflower plots in figure 3 show that subjects on average estimate their rank fairly well as most observations are close to the 45 degree line. The correlation of the rank belief in the memory test with the actual rank in the test is 0.75, whereas the correlation if the rank belief in the knowledge test with the actual rank in this test is 0.60. The regression lines in both plots are largely below the 45 degree line indicating that participants on average slightly overestimate their relative performance in both tests (by 0.7 and 1.5 ranks in the memory test and the knowledge test, respectively).[21]

[21] With respect to unincentivized estimates elicited before the feedback intervention, the correlation of beliefs in own performance with one's actual performance (in correct answers) were 0.53 and 0.18 for the memory and the knowledge test, respectively. The correlation between beliefs about one's group‘s average performance and one's group's actual average performance is 0.10 with respect to the memory test and -0.03 with respect to the knowledge test. Thus, uncertainty was generally high before the intervention, particularly so with respect to other's performance and the prior knowledge dimension. Participants before the intervention were on average slightly underconfident with respect to their own performance (by 0.2 and 1.5 points in the memory and the knowledge test, respectively) and slightly overestimated their group's average performance (by 0.7 and 0.12 points, respectively).

Figure 4: Association of Confidence in Learning Ability and in Prior Knowledge with Investment in Learning



Note: This figure shows quadratic predictions of learning investment as a function of confidence in learning ability and confidence in prior knowledge.

Figure 4 shows quadratic predictions of investment behavior as a function of the respective belief measured in percentile ranks. To facilitate interpretation of coefficients we computed inverted rank beliefs and standardized them to percentile ranks such that the maximum possible level of confidence is 100 and the minimum possible level of confidence is 0.

They thus show the quadratic approximation of the expectation about the level of investment conditional on the two confidence dimensions. As can be seen in the left panel of figure 4, there is a monotonically increasing relationship between confidence in learning ability and monetary investments in learning. The better a person thinks her memory is compared to other people, the larger the amount of costly information she acquires for the study period. The right panel of figure 4 shows that the relationship between the belief in level of prior knowledge and the investment in studying is hump shaped. Investment seems to be the highest if the person thinks that her knowledge is about average.[22]

In the following we will investigate whether these correlations between beliefs and investments are indeed driven by a direct causal effect of beliefs on investments. In order to do so, we will first check whether our random feedback manipulation affects beliefs as expected. After ensuring that it does, we will use our manipulation to instrument the beliefs in instrumental variable regressions explaining behavior and outcomes. By doing so, we will only use the exogenous

[22] A fractional polynomial plot shows nearly exactly the same hump shaped pattern.

component of beliefs, uncorrelated with other unobserved individual traits, to explain behavior.

5.2 Effect of the Feedback Manipulation on Beliefs

In order to identify the effect of our feedback manipulation on participants' beliefs, we first regress our incentivized measures of confidence in learning ability and confidence in knowledge (i.e. the subjects' beliefs about their respective rank in the considered dimension elicited after the feedback, inverted and standardized to percentile ranks) on the exogenously varied noise terms. We thus estimate the following specification by ordinary least squares, which will also constitute the first stage in our instrumental variable (IV) regressions below:

$$Confidence_i = \alpha + \beta NoiseTermMemory_i + \gamma NoiseTermKnowledge_i + \delta Controls_i + \epsilon_i \quad (3)$$

In these, as well as in all of the following regressions, we include the results of the memory and the knowledge test. Additionally, we include dummies for gender, field of study, semester of study, school GPA, income and session as control variables in all regressions.[23] All regressions also include a constant.

Table 1: First Stage Regressions

	(1) Confidence Memory	(2) Confidence Knowledge
Noise Term Memory	7.620***	-0.142
	(16.30)	(-0.30)
Noise Term Knowledge	-0.442	5.853***
	(-0.93)	(12.87)
Sum Memory Test	8.564***	-0.392
	(33.77)	(-1.64)
Sum Knowledge Test	-0.453	5.886***
	(-1.40)	(17.42)
Female	-0.796	-5.040***
	(-0.55)	(-3.48)
R^2	0.767	0.625
Sample Size	615	615

Note: OLS estimates with robust standard errors; t statistics in parentheses; both regressions contain a constant; additional control variables: dummy variables for gender, field of study (10), semester of study (22), school GPA (25), income (14) and session (24); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The results are reported in table 1 and show that the respective noise term indeed has a strong effect on the participants' beliefs about their memory and their knowledge. A one unit

[23] Tables in appendix D report the regressions without these control variables.

increase in the noise term in the memory feedback on average causes participants to believe that their memory is 7.6 percentile ranks better whereas a one unit increase in the noise term in the knowledge feedback on average causes participants to believe their knowledge is 5.9 percentile ranks better. Note that both coefficients have about the same magnitude as the respective coefficients of the true outcomes of the ability tests. Hence, our manipulation worked and the exogenous variation in feedback scores indeed affects beliefs. In the following two subsections, we can now use the manipulation to study the causal effect of confidence in learning ability and prior knowledge on investment behavior and test outcomes.

While studying the association between gender and confidence is not the focus of our study and we thus did not have an ex-ante hypothesis, it is interesting to note that women are significantly less confident than men with respect to their prior knowledge (skill level) but not so with respect to their memory (ability to acquire new skills). This further hints towards the importance of a multidimensional understanding of confidence for explaining gender effects in competitive settings. [24]

5.3 Causal Effect of Beliefs on Learning Investments

By studying whether our treatment affected behavior through affecting beliefs we can address the question of whether the relationships presented in figure 4 indeed reflect causal effects. This will allow us to test the hypotheses stated in section 3. In order to do so, we run an instrumental variable regression of beliefs on investments where the two beliefs are instrumented by the two noise terms. The first stage of the IV regression is given by equation 3. As to the second stage, we start by estimating the specification

$$Investment_i = \alpha + \beta ConfidenceMemory_i + \gamma ConfidenceKnowledge_i + \delta Controls_i + \epsilon_i \quad (4)$$

on the whole sample, including our battery of control variables. Given the hump shaped prediction with respect to the effect of confidence in prior knowledge and the availability of only two instruments, we then split the sample at the median outcome of the knowledge test[25] and

[24] In settings where skill level is important, women are observed to shy away from competition, partly due to lower confidence (Niederle and Vesterlund, 2007). It should be further explored what happens in settings where beliefs about the ability to learn play a role and whether the common finding that women are less confident than men holds in these settings.

[25] Median performance was 9 correct sets and we have 119 observation exactly at the median.

estimate effects for the worse half and the better half separately. The results are reported in table 2.

Table 2: Confidence on Investment (IV)

	(1) Invest. (All) (IV)	(2) Invest. (Better) (IV)	(3) Invest. (Worse) (IV)
Confidence Memory	0.00792**	0.00949**	0.0117***
	(2.42)	(2.00)	(2.67)
Confidence Knowledge	-0.00138	-0.00871*	0.0147**
	(-0.32)	(-1.76)	(2.24)
Sum Memory Test	0.0596*	0.0212	0.0350
	(1.92)	(0.48)	(0.85)
Sum Knowledge Test	-0.0159	-0.0822*	-0.115**
	(-0.54)	(-1.94)	(-2.42)
Female	0.00669	-0.124	0.232**
	(0.08)	(-1.23)	(2.02)
R^2	0.319	0.391	0.486
Sample Size	615	353	262
F-Test (weak ID), Memory	136.6	56.55	52.08
F-Test (weak ID), Knowledge	83.17	54.33	26.71

Note: Two-stage least squares estimates with robust standard errors; t statistics in parentheses; all regressions contain a constant; additional control variables: dummy variables for gender,field of study (10), semester of study (22), school GPA (25), income (14) and session (24); Model 1: whole sample; Model 2: performance at or above the median in knowledge test; Model 3: below median performance in knowledge test; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column (1) of table 2 shows that confidence in learning ability significantly increases investment whereas the effect of confidence in levels of prior knowledge is insignificant when looking at the whole sample. Since we expected a positive effect for individuals with low prior knowledge and a negative effect for individuals with high prior knowledge, we split the sample. In columns (2) and (3) we can see that both in the better and in the worse half of participants, confidence in learning ability has a positive effect on learning investment. In line with our predictions, we also observe that confidence in levels of knowledge has a negative effect on individuals with above average levels of prior knowledge but a positive effect on individuals with below average levels of prior knowledge. More specifically, for confidence in learning ability we find that an increase of confidence by 10 percentile ranks raises investment in learning by about 9 euro cents for the better half of students and about 12 euro cents for the worse half of students. These effects are significant at the 5 percent and the 1 percent level, respectively. For confidence in knowledge we find that an increase of confidence by 10 percentile ranks lowers investment in learning by about 9 euro cents for students with above average level of prior knowledge but raises investment in learning by about 15 euro cents for students with below average level of prior knowledge. These effects, respectively, are significant at the 10 percent and the 5 percent level. F-tests indicate that our instruments are sufficiently strong.

The experimental results show that beliefs about abilities causally affect how much a person invests in learning. We find that people on average make larger investments in learning the better they believe their learning ability to be. We also find evidence in favor of the hypothesis that increasing the confidence in prior knowledge reduces incentives for individuals whose knowledge is already above average but increases incentives for individuals whose knowledge is below average.

5.4 Causal Effect of Beliefs on Test Outcomes

We are also interested in whether the behavioral change we brought about by changing confidence beliefs has an effect on students' outcomes in the final test. We begin by estimating how beliefs causally affect the rank one received in the final test. Note that the first stage of the IV regressions is again given by equation 3. The second stage is given by:

$$Rank_i = \alpha + \beta ConfidenceMemory_i + \gamma ConfidenceKnowledge_i + \delta Controls_i + \epsilon_i \quad (5)$$

As can be seen in columns (1) and (2) of table 3, for the better half of participants in the knowledge test we find no effect of confidence in learning ability[26] but we do find a negative effect of confidence in knowledge again. As confidence in knowledge increases by one percentile rank the outcome in the final test decreases by about 0.3 percentile ranks. For the worse half of participants in the knowledge test we find that as confidence in learning ability increases by one percentile rank the outcome in the final test increases by about 0.3 percentile ranks, while as confidence in prior knowledge increases by one percentile rank the outcome in the final test increases by about 0.5 percentile ranks.

We then use an IV probit estimation method based on Newey (1987) to test whether beliefs also causally affect the probability of passing the test. The first stage is again given by equation 3. The second stage is a probit regression of the form

[26]Hence, for subjects in the better half, the effect of a higher confidence in learning ability on higher learning investments does not translate into better test outcomes. One possible explanation is a physical limitation to the subjects' short term working memory. While more confident subjects were further motivated to acquire knowledge (and thus invested more), they may have been unable to memorize this information appropriately in the given time frame.

$$Pr(y=1|x) = G(\alpha + \beta ConfidenceMemory_i + \gamma ConfidenceKnowledge_i + \delta Controls + \epsilon_i) \quad (6)$$

As can be seen by looking at columns (3) and (4) of table 3, and analogously to the results in columns (1) and (2), we find that raising the confidence in memory increases the passing probability of people who performed in the worse half in the knowledge test, whereas raising the confidence in prior knowledge decreases the passing probability of above average and increases the passing probability of below average performers in the knowledge test. We do not find a significant effect of confidence in memory for individuals who performed in the better half in the knowledge test.

Table 3: Confidence on Rank and Probability of Passing Final Test (IV)

	(1) Rank (Better)	(2) Rank (Worse)	(3) Pr. Pass. (Better)	(4) Pr. Pass. (Worse)
Confidence Memory	-0.108	0.320**	-0.00106	0.0267**
	(-0.66)	(2.26)	(-0.12)	(2.13)
Confidence Knowledge	-0.297*	0.549**	-0.0190*	0.0398**
	(-1.68)	(2.36)	(-1.93)	(2.23)
Sum Memory Test	3.478**	0.722	0.111	-0.0846
	(2.23)	(0.53)	(1.30)	(-0.71)
Sum Knowledge Test	1.071	-2.137	0.0458	-0.120
	(0.71)	(-1.28)	(0.53)	(-0.92)
Female	-10.13***	0.496	-0.502***	-0.299
	(-3.02)	(0.14)	(-2.64)	(-0.94)
R^2	0.234	0.375		
Sample Size	353	262	339	235
F-Test (weak ID), M.	56.55	52.08		
F-Test (weak ID), K.	54.33	26.71		

Note: Models 1 and 2: two-stage least squares estimates with robust standard errors; Models 3 and 4: Newey's two-step estimator for binary endogenous variables; t statistics in parentheses; all regressions contain a constant; additional control variables: dummy variables for gender,field of study (10), semester of study (22), school GPA (25), income (14) and session (24); Models 1 and 3: performance at or above the median in knowledge test; Model 2 and 4: below median performance in knowledge test; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Conclusion

We studied the causal effects of confidence in prior knowledge and in the ability to learn in a lab experimentto. Based on a simple formal model, we hypothesized that a higher confidence in one's level of prior knowledge causes students with low levels of knowledge to invest more. This is because it raises the probability that an additional remembered fact is pivotal to passing the test. For students with high levels of prior knowledge we expected the opposite, i.e. that raising their confidence in knowledge would lower their effort to prepare for the test because it subjectively moves them further away from the passing threshold such that learning becomes less relevant for whether someone passes or fails the test. For the other dimension, confidence in one's learning ability, we expected that raising this dimension of confidence would cause students to invest more effort in learning irrespective of the prior knowledge because the perceived marginal cost of effort decreases.

Our results support these hypotheses. Confidence in learning ability, indeed, raises learning investments irrespective of the prior level of knowledge, whereas confidence in prior knowledge has a negative effect on individuals with above average prior knowledge and a positive effect on individuals with below average prior knowledge on investments. Some of the behavioral effects of our feedback intervention are also reflected by the test outcomes. Raising confidence in learning ability improves the rank and increases the probability of an individual with below average prior knowledge passing the test, whereas we do not find a significant effect for the rank or passing probability of above average individuals. Furthermore, raising confidence in prior knowledge improves the rank and increases the probability that an individual with below average prior knowledge passes the test, whereas it worsens the rank and decreases the passing probability of individuals with above average prior knowledge.

We thus have shown that confidence affects investments in learning in very different ways depending on the specific dimension the belief refers to. People invest more in learning when their confidence in the ability to learn is raised and we find no evidence of a detrimental effect of "too much confidence" in learning ability. Of course, we caution that we studied a lab experiment in a specific content area, and further work has to be done to investigate the validity of the results in other contexts. Our results that different dimensions of motivational beliefs exhibit different functional relationships with effort and outcomes show that generalized statements about the role of confidence in competitive settings can be misleading and confidence should be viewed as a multidimensional concept. For instance, our observation that there are gender differences in confidence in prior knowledge, as has often been found in the past, but not with respect to confidence in learning ability suggests that the multidimensionality of ability beliefs might be very relevant for explaining why different behavior of men and women can be observed in some settings but not in others.

Insights about the different effects of confidence in learning ability and confidence in prior knowledge have implications not only for the design of interventions aimed at positively affecting academic motivation but also for subjective performance evaluation policies in firms and other organizations. A large literature in psychology and economics has, for instance, stressed that subjective performance evaluations tend to be biased and, in particular, evaluators often tend to be too lenient (see e.g. Murphy and Cleveland 1995; Prendergast 1999). Our results imply that rater leniency (i.e. the tendency to assign too generous performance ratings) can raise motivation when the rater assesses an individual's ability to learn. However, leniency in the rating of a skill level can reduce the motivation as it may signal that one has "already done enough". Hence, while raising confidence in the ability to acquire a certain skill or achieve an outcome can be beneficial, raising confidence in the skill itself or the level of past achievements can be detrimental.

Finally, we note that while we wanted to identify the causal effect of confidence on performance, we did not intend to evaluate the usefulness of confidence *manipulations* in real world settings. The confidence manipulation through noise terms added to test results is designed as a research tool that makes it possible to study causal effects of confidence. It is not meant as an intervention that should be implemented to raise confidence in field settings but we believe that our work can inform the optimal design of interventions that aim at influencing confidence to raise motivation in the field. For instance, our results indicate that interventions that raise the confidence in the ability to learn and grow should be beneficial. Our results are thus well in line with the idea of inducing a "growth mindset", i.e. the belief that intelligence is malleable rather than fixed, which has been shown to raise educational outcomes (Yeager et al., 2014; Paunesku et al., 2015; Alan et al., 2016). However, our results also show that interventions that raise confidence in traits that directly contribute to outcomes (such a prior knowledge) may be detrimental.

References

Ackerman, P. L., Beier, M. E., and Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs?. *Psychological Bulletin*, 131(1):30–60.

Alan, S., Boneva, T., and Ertac, S. (2016). Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *HCEO Working Paper*.

Alloway, T. P. and Alloway, R. G. (2010). Investigating the predictive roles of working memory and iq in academic attainment. *Journal of Experimental Child Psychology*, 106(1):20–29.

Aoyagi, M. (2010). Information feedback in a dynamic tournament. *Games and Economic Behavior*, 70(2):242–260.

Azmat, G., Bagues, M., Cabrales, A., and Iriberri, N. (2015). What you know can't hurt you (for long): A field experiment on relative performance feedback in higher education. *unpublished working paper*.

Azmat, G. and Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94(7-8):435–452.

Bandiera, O., Larcinese, V., and Rasul, I. (2015). Blissful ignorance? a natural experiment on the effect of feedback on students' performance. *Labour Economics*, 34:13–25.

Benabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.

Benabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70(3):489–520.

Caliendo, M., Cobb-Clark, D. A., and Uhlendorff, A. (2015). Locus of control and job search strategies. *Review of Economics and Statistics*, 97(1):88–103.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Casas-Arce, P. and Martínez-Jerez, F. A. (2009). Relative performance compensation, contests, and dynamic incentives. *Management Science*, 55(8):1306–1320.

Cebi, M. (2007). Locus of control and human capital investment revisited. *The Journal of Human Resources*, 42(4):919–932.

Compte, O. and Postlewaite, A. (2004). Confidence-enhanced performance. *American Economic Review*, 94(5):1536–1557.

Costa-Gomes, M. A., Huck, S., and Weizsaecker, G. (2014). Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect. *Games and Economic Behavior*, 88(0):298–309.

Dechenaux, E., Kovenock, D., and Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4):609–669.

Eccles, J. S. and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1):109–132.

Ederer, F. (2010). Feedback and motivation in dynamic tournaments. *Journal of Economics & Management Strategy*, 19(3):733–769.

Filippin, A. and Paccagnella, M. (2012). Family background, self-confidence and economic outcomes. *Economics of Education Review*, 31(5):824–834.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Gershkov, A. and Perry, M. (2009). Tournaments with midterm reviews. *Games and Economic Behavior*, 66(1):162–190.

Greiner, B. (2004). An online recruitment system for economic experiments. In Kremer, K. and Macho, V., editors, *Forschung und wissenschaftliches Rechnen*, pages 79–93. Ges. fuer Wiss. Datenverarbeitung, Goettingen.

Grossman, Z. and Owens, D. (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior & Organization*, 84(2):510–524.

Hattie, J. and Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1):81–112.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *IZA Discussion Papers*, 6580.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Heineck, G. and Anger, S. (2010). The returns to cognitive abilities and personality traits in germany. *Labour Economics*, 17(3):535–546.

Kluger, A. N. and DeNisi, A. (1998). Feedback interventions: Toward the understanding of a double-edged sword. *Current Directions in Psychological Science*, 7(3):67–72.

Koch, A., Nafziger, J., and Nielsen, H. S. (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization*, 115:3–17.

Kuhnen, C. M. and Tymula, A. (2012). Feedback, self-esteem, and performance in organizations. *Management Science*, 58(1):94–113.

Lazear, E. P. and Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, 89(5):841–864.

Mobius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2011). Managing self-confidence: Theory and experimental evidence. *NBER Working Papers*, 17014.

Murphy, K. R. and Cleveland, J. N. (1995). *Understanding Performance Appraisal*. Sage, Thousand Oaks.

Newey, W. K. (1987). Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics*, 36(3):231–250.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

Paunesku, D., Walton, G., Romero, C., Smith, E., Yeager, D., and Dweck, C. (2015). Mind-Set Interventions Are a Scalable Treatment for Academic Underachievement. *Psychological Science*, 26(6):784–793.

Prendergast, C. J. (1999). The provision of incentives in firms. *Journal of Economic Literature*, 37:7–63.

Reuben, E., Wiswall, M., and Zafar, B. (2013). Preferences and biases in educational choices and labor market expectations: shrinking the black box of gender. *Federal Reserve Bank of New York Staff Reports*, 627.

Santos-Pinto, L. (2008). Positive self-image and incentives in organisations. *The Economic Journal*, 118(531):1315–1332.

Schwardmann, P. and Van der Weele, J. J. (2016). Deception and self-deception. *Tinbergen Institute Discussion Papers*, 2016-12.

Spinnewijn, J. (2015). Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association*, 13(1):130–167.

Tran, A. and Zeckhauser, R. (2012). Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics*, 96(9-10):645–650.

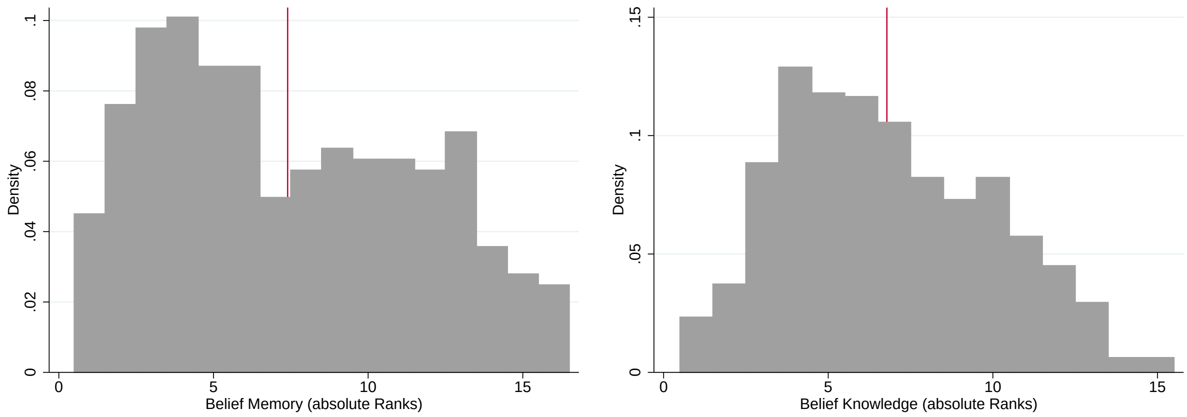
Wilhelm, O., Hildebrandt, A. H., and Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, 4(433).

Yeager, D. S., Johnson, R., Spitzer, B. J., Trzesniewski, K. H., Powers, J., and Dweck, C. S. (2014). The far-reaching effects of believing people can change: Implicit theories of personality shape stress, health, and achievement during adolescence. *Journal of Personality and Social Psychology*, 106(6):867 – 884.

Appendix

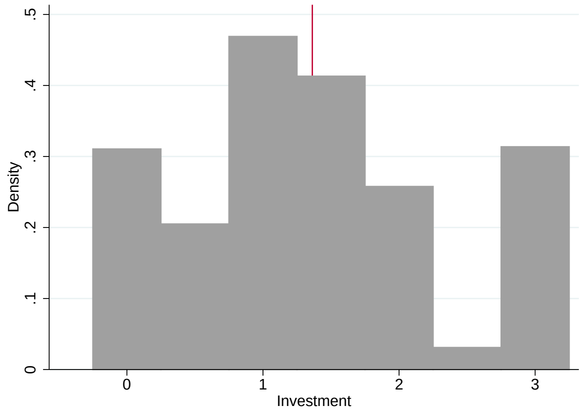
A Descriptive Statistics and Figures

Figure A.1: Rank Beliefs



Note: Distributions and means of rank beliefs elicited after giving feedback. (1 is best)

Figure A.2: Investment (in Euros)



Note: Distribution of learning investments in euros.

Table A.1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Noise Term Memory	-0.03	1.42	-2	2	644
Noise Term Knowledge	0.01	1.4	-2	2	644
Belief Memory	54	28.54	6.25	100	644
Belief Knowledge	58.32	21.51	7.69	100	644
Sum Memory Test	5.15	2.55	0	11	644
Sum Knowledge Test	8.87	2.18	0	16	644
Sum Test 3	10.72	2.52	1	20	644
Investment	1.36	0.95	0	3	644
Prob. of Passing Test 3	0.5	0.5	0	1	644
Profit	11.29	5.02	3.2	19.4	644
Female	0.63	0.48	0	1	644
School GPA	2.05	0.6	1	3.5	623
Humanities	0.16	0.37	0	1	644
Social Sciences	0.09	0.29	0	1	644
Law	0.05	0.22	0	1	644
Business	0.26	0.44	0	1	644
Economics	0.13	0.34	0	1	644
Medicine	0.05	0.21	0	1	644
Natural Sciences	0.08	0.27	0	1	644
Psychology	0.01	0.12	0	1	644
Other Subjects	0.14	0.35	0	1	644
Non-Student	0.02	0.13	0	1	644
Semester 1	0.06	0.24	0	1	635
Semester 2	0.11	0.31	0	1	635
Semester 3	0.06	0.23	0	1	635
Semester 4	0.12	0.33	0	1	635
Semester 5	0.08	0.27	0	1	635
Semester 6	0.13	0.33	0	1	635
Semester 7	0.08	0.27	0	1	635
Semester 8	0.1	0.31	0	1	635
Semester 9	0.06	0.24	0	1	635
Semester 10	0.05	0.22	0	1	635
Semester 11	0.04	0.2	0	1	635
Semester 12	0.03	0.18	0	1	635
Semester 13	0.02	0.14	0	1	635
Semester 14	0.01	0.11	0	1	635
Semester 15	0.01	0.1	0	1	635
Semester 16	0.01	0.1	0	1	635
Semester 17	0	0.06	0	1	635
Semester 18	0.01	0.08	0	1	635
Semester 19	0	0.04	0	1	635
Semester 20	0	0.04	0	1	635
Semester 21	0	0.06	0	1	635
Semester 23	0	0.04	0	1	635
Session 1	0.04	0.19	0	1	644
Session 2	0.04	0.19	0	1	644
Session 3	0.04	0.2	0	1	644
Session 4	0.03	0.17	0	1	644
Session 5	0.04	0.19	0	1	644
Session 6	0.05	0.21	0	1	644
Session 7	0.03	0.18	0	1	644
Session 8	0.04	0.19	0	1	644
Session 9	0.05	0.21	0	1	644
Session 10	0.05	0.21	0	1	644
Session 11	0.05	0.22	0	1	644
Session 12	0.03	0.17	0	1	644
Session 13	0.05	0.22	0	1	644
Session 14	0.04	0.2	0	1	644
Session 15	0.03	0.16	0	1	644
Session 16	0.03	0.18	0	1	644
Session 17	0.05	0.21	0	1	644
Session 18	0.05	0.22	0	1	644
Session 19	0.05	0.22	0	1	644
Session 20	0.05	0.21	0	1	644
Session 21	0.04	0.2	0	1	644
Session 22	0.05	0.22	0	1	644
Session 23	0.04	0.2	0	1	644
Session 24	0.05	0.21	0	1	644

B OLS Regressions of Beliefs on Behavior and Outcomes

Table B.1: Confidence on Investment (OLS)

	(1) Invest. (All)	(2) Invest. (Better)	(3) Invest. (Worse)
Belief Memory	0.00871***	0.00637**	0.0120***
	(3.97)	(2.03)	(3.38)
Belief Knowledge	0.00239	0.00234	0.000527
	(1.11)	(0.78)	(0.15)
Sum Memory Test	0.0552**	0.0550	0.0189
	(2.15)	(1.53)	(0.47)
Sum Knowledge Test	-0.0380*	-0.146***	-0.0379
	(-1.84)	(-3.47)	(-0.89)
Female	0.0264	-0.0559	0.205
	(0.32)	(-0.50)	(1.51)
R^2	0.323	0.421	0.528
Sample Size	615	353	262

Note: OLS estimates with robust standard errors; t statistics in parentheses; all regressions contain a constant; additional control variables: dummy variables for gender, field of study (10), semester of study (22), school GPA (25), income (14) and session (24); * p<0.10, ** p<0.05, *** p<0.01.

Table B.2: Confidence on Outcomes (OLS)

	(1) Rank (Better)	(2) Rank (Worse)	(3) Prob. Pass. (Better)	(4) Prob. Pass. (Worse)
Belief Memory	-0.0501	0.103	-0.0000928	0.00961
	(-0.50)	(1.00)	(-0.02)	(1.51)
Belief Knowledge	-0.0158	0.283**	-0.00992*	0.0210***
	(-0.14)	(2.15)	(-1.91)	(2.79)
Sum Memory Test	3.149***	2.368**	0.107**	0.0472
	(2.83)	(2.09)	(2.03)	(0.70)
Sum Knowledge Test	-0.481	-1.090	-0.00528	-0.0569
	(-0.32)	(-0.65)	(-0.08)	(-0.66)
Female	-8.536**	0.786	-0.446**	-0.222
	(-2.23)	(0.19)	(-2.45)	(-0.87)
R^2	0.254	0.405		
Chi^2			84.89	137.9
Sample Size	353	262	339	235

Note: OLS estimates with robust standard errors; t statistics in parentheses; all regressions contain a constant; additional control variables: dummy variables for gender, field of study (10), semester of study (22), school GPA (25), income (14) and session (24); * p<0.10, ** p<0.05, *** p<0.01.

C Reduced Form Estimates

Table C.1: Noise Terms on Investment (OLS)

	(1) Invest. (All) (OLS)	(2) Invest. (Better) (OLS)	(3) Invest. (Worse)(OLS)
Noise Term Memory	0.0606**	0.0660*	0.0917**
	(2.21)	(1.74)	(2.01)
Noise Term Knowledge	-0.0116	-0.0534	0.0797*
	(-0.42)	(-1.53)	(1.84)
Sum Memory Test	0.128***	0.103***	0.125***
	(8.35)	(4.90)	(4.96)
Sum Knowledge Test	-0.0276*	-0.136***	-0.0572
	(-1.65)	(-3.65)	(-1.34)
Female	0.00732	-0.0847	0.231*
	(0.09)	(-0.77)	(1.71)
R^2	0.305	0.423	0.506
Sample Size	615	353	262

Note: OLS estimates with robust standard errors; t statistics in parentheses; all regressions contain a constant; additional control variables: dummy variables for gender, field of study (10), semester of study (22), school GPA (25), income (14) and session (24); * p<0.10, ** p<0.05, *** p<0.01.

Table C.2: Noise Terms on Outcomes (OLS)

	(1) Rank (Better) (OLS)	(2) Rank (Worse)(OLS)	(3) Prob. Winning (Better) (Probit)	(4) Prob. Winning (Worse)(Probit)
main				
Noise Term Memory	-0.703	2.408*	-0.00914	0.182**
	(-0.55)	(1.66)	(-0.15)	(2.17)
Noise Term Knowledge	-1.861	2.979**	-0.118*	0.219***
	(-1.46)	(1.98)	(-1.96)	(2.77)
Sum Memory Test	2.679***	3.065***	0.109***	0.118***
	(3.74)	(4.08)	(3.16)	(2.92)
Sum Knowledge Test	-0.587	0.168	-0.0589	0.0237
	(-0.44)	(0.10)	(-0.94)	(0.29)
Female	-8.407**	0.348	-0.401**	-0.302
	(-2.19)	(0.08)	(-2.19)	(-1.21)
R^2	0.261	0.404		
Chi^2			83.82	133.6
Sample Size	353	262	339	235

Note: OLS estimates with robust standard errors; t statistics in parentheses; all regressions contain a constant; additional control variables: dummy variables for gender, field of study (10), semester of study (22), school GPA (25), income (14) and session (24); * p<0.10, ** p<0.05, *** p<0.01.

D Results Without Session Dummies and Demographic Control Variables

Table D.1: First Stage Regressions Without Additional Control Variables

	(1) Confidence Memory	(2) Confidence Knowledge
Noise Term Memory	7.468***	-0.582
	(18.04)	(-1.39)
Noise Term Knowledge	-0.684*	6.239***
	(-1.70)	(14.96)
Sum Memory Test	8.551***	-0.204
	(34.89)	(-0.87)
Sum Knowledge Test	-0.412	5.909***
	(-1.40)	(19.71)
Constant	13.88***	6.864**
	(4.50)	(2.36)
R^2	0.727	0.530
Sample Size	644	644

OLS estimates with robust standard errors; t statistics in parentheses;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.2: Confidence on Investment (IV) Without Additional Control Variables

	(1) Invest. (All)	(2) Invest. (Better)	(3) Invest. (Worse)
Confidence Memory	0.00906***	0.00646	0.0118**
	(2.67)	(1.34)	(2.42)
Confidence Knowledge	-0.000832	-0.00581	0.00627
	(-0.21)	(-1.14)	(0.91)
Sum Memory Test	0.0531*	0.0718	0.0373
	(1.68)	(1.61)	(0.83)
Sum Knowledge Test	-0.00194	-0.0467	-0.0662
	(-0.07)	(-1.07)	(-1.37)
Constant	0.666***	1.557***	0.654**
	(4.05)	(3.98)	(2.16)
R^2	0.164	0.137	0.175
Sample Size	644	374	270
F-Test (weak ID), Memory	164.4	87.01	80.77
F-Test (weak ID), Knowledge	112.4	78.50	36.53

Two-stage least squares estimates with robust standard errors; t statistics in parentheses; Model 1: whole sample; Model 2: performance at or above the median in knowledge test; Model 3: below median performance in knowledge test; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.3: Confidence on Rank and Probability of Passing Final Test (IV) Without Additional Control Variables

	(1) Rank (Better)	(2) Rank (Worse)	(3) Pr. Pass. (Better)	(4) Pr. Pass. (Worse)
Confidence Memory	-0.114	0.258*	-0.00288	0.00704
	(-0.76)	(1.74)	(-0.42)	(0.99)
Confidence Knowledge	-0.199	0.370*	-0.0128*	0.0151
	(-1.25)	(1.69)	(-1.71)	(1.49)
Sum Memory Test	3.750***	1.024	0.126*	0.0297
	(2.70)	(0.72)	(1.93)	(0.45)
Sum Knowledge Test	0.973	-0.794	0.0215	-0.0573
	(0.68)	(-0.50)	(0.34)	(-0.77)
Constant	51.81***	24.15**	0.261	-0.989**
	(3.85)	(2.24)	(0.45)	(-2.02)
R^2	0.0334	0.0756		
Sample Size	374	270	374	270
F-Test (weak ID), M.	87.01	80.77		
F-Test (weak ID), K.	78.50	36.53		

Two-stage least squares estimates with robust standard errors; t statistics in parentheses; Model 1: whole sample; Model 2: performance at or above the median in knowledge test; Model 3: below median performance in knowledge test; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E Timeline of the Experiment

1. Measurement of Prior Knowledge and Learning Ability: Subjects take two tests (incentivized with piece rate, the order is randomized to control for ordering effects):

    - „knowledge test“: participants have to solve 20 sets of three German cities each by indicating which is the largest, which is the second largest and which is the third largest in terms of population within each triple
    - „memory test“: participants for 15 minutes see a screen with a list of 36 German cities with (arbitrary) four digit „cities codes“ which they can memorize, then they have to solve 12 sets of three cities each by indicating which one has the largest, which one has the second largest, and which one has the third largest city code
    - Immediately after each test participants estimate their number of correct sets and other‘s average number of correct sets in each test (belief elicitation, unincentivized)

2. Information on further course (introduction of combined test): Subjects are informed that there will be a third test and that they earn a prize if their outcome is above average. They are explained how they can prepare for it. Furthermore, they are told that

they will receive feedback and given an explanation of how the feedback is computed.

3. Feedback stage: Subjects receive noisy feedback about their performance in both tests (treatment variation)

4. Measurement of confidence (belief elicitation, both tests, incentivized): Subjects estimate their rank in both tests

5. Investment stage (information acquisition): Subjects receive a budget of 3 euros from which they can buy information on cities in increments of 0.5 euros or 10 cities (behavioral outcome variable)

6. Measurement of outcomes (combined knowledge and memory test): Subjects take the third test (economic outcome variables)

F Details on the Tests, Feedback, Elicitation of Beliefs, and Investment Stage

The experiment was conducted in German, so in the following we give the English translation of the texts. All the cities used in the experiment come from the set of the 200 largest cities in Germany. We pretested all instructions and tests to ensure that they are understandable and produced a sufficient variance of results so that relative performance/ability could be measured precisely. Before the tests started, an introductory screen described the test and how money could be earned. We also made sure that subjects understood the rules of the tests by including a sample exercise before each test and subjects could only start the test after answering it according to the rules.

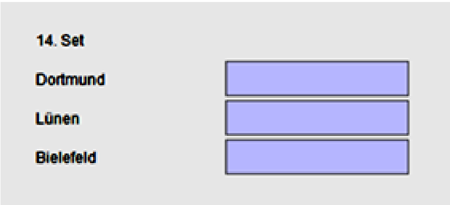
F.1 Description of Knowledge Test

The instruction on the introductory screen to the knowledge test said:

> "In the following you can earn money by ordering three cities, respectively, according to their number of inhabitants. In total there are 20 sets of 3 cities each. For each completely correct set you will receive 0.10 euros. If the set was not answered completely correctly you will not receive any money for it. You have 6 minutes to work on the test. Write a 1 in the field next to the city you belief is

the largest of the three, write a 2 in the field of the intermediate city and write a 3 in the field next to the smallest city."

On the test screen itself a summary of the instructions and the payment scheme was given. A countdown clock was shown. For example, a set of three cities looked like this:



F.2 Description of Memory Test

The instruction on the introductory screen to the memory test said:

> "In the following you can earn money by ordering three cities, respectively, according to their city codes. In total there are 12 sets of 3 cities each. For each completely correct set you will receive 0.20 euros. If the set was not answered completely correctly you will not receive any money for it. You have 6 minutes to work on the test.
>
> Since the city codes are generally not known, you will receive an alphabetically ordered list with all 36 cities and their respective city codes. This list will be displayed to you in a learning phase of 15 minutes. You have the opportunity to memorize the ranking (relative size) of these city codes, in order to later order three cities each according to this number. During the test this list will not be displayed anymore, so that only your memory will help you to do the ordering. Note-taking is not allowed.Violation of this rule will lead to the exclusion from this and future experiments.
>
> Write a 1 in the field next to the city which according to your memory has the largest city code, write a 2 in the field of the city with the second largest city code and write a 3 in the field next to the city with the smallest city code."

On the learning and test screens a summary of the instructions and the payment scheme was given. A countdown clock was shown. The sets of three cities in the memory test looked the same as in the knowledge test but none of the city names were used twice. Information displayed in the learning phase looked like this:

Friedrichshafen	5016
Görlitz	6110
Greifswald	5039
Gummersbach	4012
Hameln	2006
Heidenheim	5019
Herzogenrath	4016
Hürth	2028
Langenfeld	8020
Langenhagen	1010
Lörrach	6050
Melle	9024

F.3 Description of Feedback

After subjects have been told that there will be a third "main test" and that they can prepare for it, they are informed that they are about to receive feedback. Next, they are shown a screen where the computation of the "feedback scores" is explained:

> "The experimental software will now generate a knowledge score and a memory score for each participant. The knowledge score is being computed based on a participant's number of correct answers in the knowledge test whereas the memory score is computed based on a participant's number of correct answers in the in the memory test. In expectation, each score is equal to the participant's actual number of correct answers. The experimental software will soon let you know your score.
>
> **Computation of the feedback scores**:
>
> Your scores are composed of the following:
>
> Knowledge score = number of your correct sets in the knowledge test + random variable X

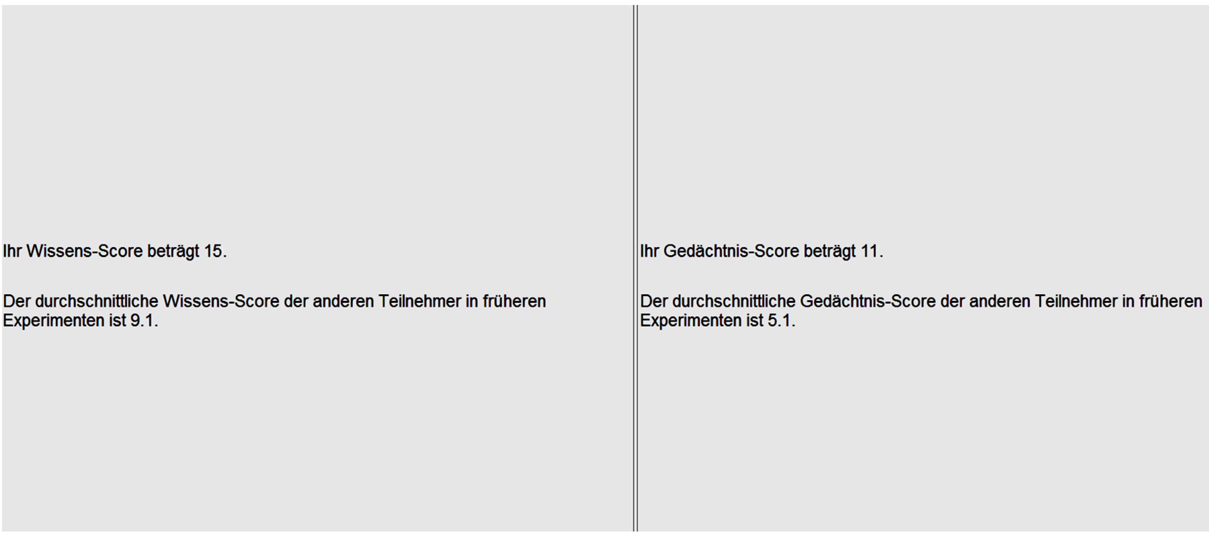
Memory Score = number of your correct sets in the memory test + random variable Y

The random variables X and Y can each assume values between -2 and +2, that means each of the values (-2, -1, 0, +1, +2) is equally likely (i.e. occurs with a probability of 20%). Furthermore, the random variables X and Y are independent of each other, that means also all combinations of values of the random variables X and Y are equally likely."

On the Next screen, subjects receive the following information:

"The knowledge score can help you to assess your knowledge of cities relative to other participants whereas the memory score can help you to assess your memory capacity relative to other participants. The two scores give your number of correct sets in each test with a certain imprecision but in expectation equal the actual number of your correct answers."

The feedback screen displayed both a participant's two scores and the respective average score of participants in earlier experimental sessions: "Your [knowledge/memory] score is [x]. The average [knowledge/memory] score of the other participants in earlier experiments is [9.1/5.1]" It looked like this:



F.4 Elicitation of Beliefs

The elicitation screen contained the following text:

"Half of participants in this room worked on the two tests in the same order as you. How do you assess your own results in both tests relative to theses participants? Please estimate your rank below. For each estimate you will earn one euro if you guess the rank exactly right. There are [x] participants in your group.

The participant with the highest number of points occupies rank 1, the participant with the lowest number of points occupies rank [x]."

Then participants could indicate their rank beliefs in the knowledge and the memory test by selecting a number on two lines of radio buttons. The number of radio buttons was automatically adjusted to the number of people in each of the two groups per session.

F.5 Investment Stage

The decision screen contained the following information:

"**Description of test 3: combined knowledge and memory test**

In the following you can earn money by ordering three cities, respectively, according to their numbers of inhabitants. In total there are 20 sets of 3 cities each. You have 6 minutes to work on the test.

The cities are German cities of comparable size and prominence as the cities in the knowledge test about the numbers of inhabitants. However, no of these cities will be in the test again.

If your result is above average, that is if you get more correct answers than the average of the participants in the room who worked on the first two tests in the same order as you, you will receive 10 euros, if not you will receive zero euros.

You have the possibility to improve you knowledge of the cities in a learning phase.

Description of preparation for test 3

In order to prepare for test 3, you may buy information about cities' numbers of inhabitants. In order to do so you receive, independently of your performance until now, a budget of 3.00 euros. The part of the budget that you do not spend, will be added to your payoff in the end of the experiment. All cities you can buy

> are part of the test. You can buy packages of 10 cities each. Each package allows you to completely answer at least 3 assignments (sets).
>
> Example for information you can buy:
>
> Innsbruck 121,329
>
> Following your selection, for 15 minutes the program will show in alphabetical order your acquired packages of cities with their respective numbers of inhabitants. This information you may memorize so that you can better order cities according to their size in the main test. Note-taking is not allowed. Violation of this rule will lead to the exclusion from this and future experiments."

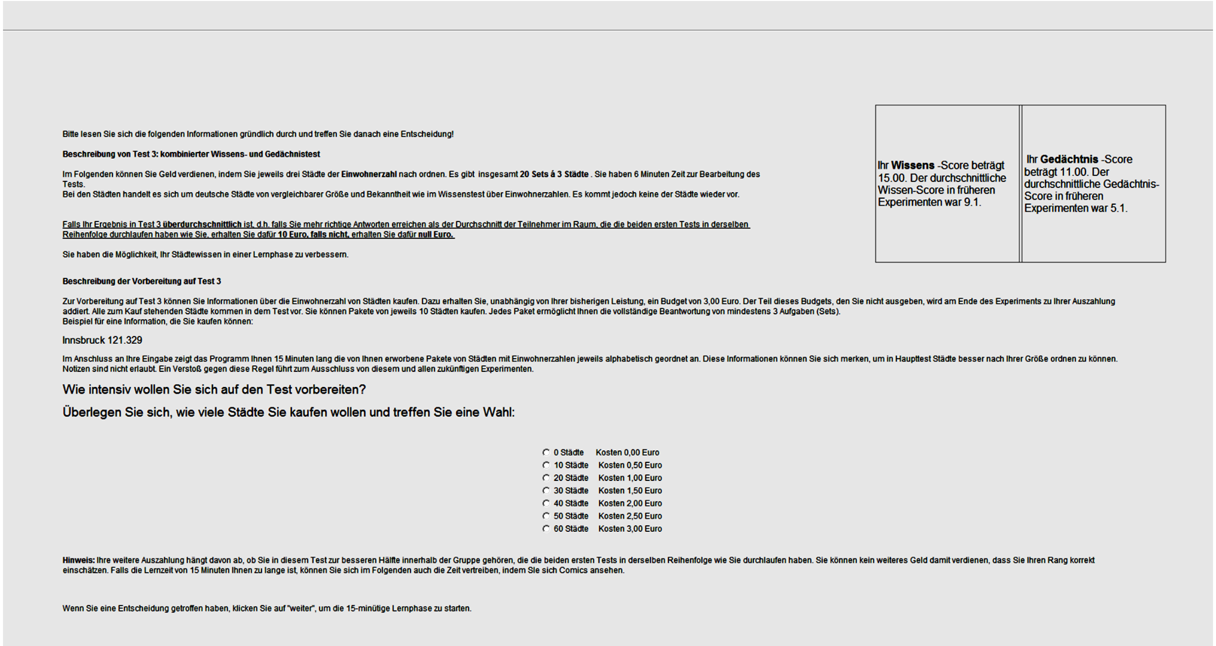
Below this text, subjects were asked to decide how many cities they want to buy and indicate their choice with the respective radio button. They have to make a choice between buying 0, 10, 20, 30, 40, 50, or 60 cities. Each ten cities cost 0.5 euros.

Below the radio buttons it said:

> "Please note: Your further payoff depends on whether you belong to the better half of the group who worked on the first two tests in the same order as you. You cannot earn additional money by estimating your rank correctly. In case you find the study time of 15 minutes too long, you can also spend time looking at comics."

A reminder of their knowledge and memory score is displayed in the upper right corner of the screen.

This is how the screen looked like:



Bitte lesen Sie sich die folgenden Informationen gründlich durch und treffen Sie danach eine Entscheidung!

Beschreibung von Test 3: kombinierter Wissens- und Gedächnistest

Im Folgenden können Sie Geld verdienen, indem Sie jeweils drei Städte der **Einwohnerzahl** nach ordnen. Es gibt insgesamt **20 Sets á 3 Städte** . Sie haben 6 Minuten Zeit zur Bearbeitung des Tests.
Bei den Städten handelt es sich um deutsche Städte von vergleichbarer Größe und Bekanntheit wie im Wissenstest über Einwohnerzahlen. Es kommt jedoch keine der Städte wieder vor.

Falls Ihr Ergebnis in Test 3 **überdurchschnittlich** ist, d.h. falls Sie mehr richtige Antworten erreichen als der Durchschnitt der Teilnehmer im Raum, die die beiden ersten Tests in derselben Reihenfolge durchlaufen haben wie Sie, erhalten Sie dafür **10 Euro, falls nicht,** erhalten Sie dafür **null Euro.**

Sie haben die Möglichkeit, Ihr Städtewissen in einer Lernphase zu verbessern.

Beschreibung der Vorbereitung auf Test 3

Zur Vorbereitung auf Test 3 können Sie Informationen über die Einwohnerzahl von Städten kaufen. Dazu erhalten Sie, unabhängig von Ihrer bisherigen Leistung, ein Budget von 3,00 Euro. Der Teil dieses Budgets, den Sie nicht ausgeben, wird am Ende des Experiments zu Ihrer Auszahlung addiert. Alle zum Kauf stehenden Städte kommen in dem Test vor. Sie können Pakete von jeweils 10 Städten kaufen. Jedes Paket ermöglicht Ihnen die vollständige Beantwortung von mindestens 3 Aufgaben (Sets).
Beispiel für eine Information, die Sie kaufen können:

Innsbruck 121.329

Im Anschluss an Ihre Eingabe zeigt das Programm Ihnen 15 Minuten lang die von Ihnen erworbene Pakete von Städten mit Einwohnerzahlen jeweils alphabetisch geordnet an. Diese Informationen können Sie sich merken, um in Haupttest Städte besser nach Ihrer Größe ordnen zu können. Notizen sind nicht erlaubt. Ein Verstoß gegen diese Regel führt zum Ausschluss von diesem und allen zukünftigen Experimenten.

Wie intensiv wollen Sie sich auf den Test vorbereiten?

Überlegen Sie sich, wie viele Städte Sie kaufen wollen und treffen Sie eine Wahl:

- 0 Städte Kosten 0,00 Euro
- 10 Städte Kosten 0,50 Euro
- 20 Städte Kosten 1,00 Euro
- 30 Städte Kosten 1,50 Euro
- 40 Städte Kosten 2,00 Euro
- 50 Städte Kosten 2,50 Euro
- 60 Städte Kosten 3,00 Euro

Hinweis: Ihre weitere Auszahlung hängt davon ab, ob Sie in diesem Test zur besseren Hälfte innerhalb der Gruppe gehören, die die beiden ersten Tests in derselben Reihenfolge wie Sie durchlaufen haben. Sie können kein weiteres Geld damit verdienen, dass Sie Ihren Rang korrekt einschätzen. Falls die Lernzeit von 15 Minuten Ihnen zu lange ist, können Sie sich im Folgenden auch die Zeit vertreiben, indem SIe sich Comics ansehen.

Wenn Sie eine Entscheidung getroffen haben, klicken Sie auf "weiter", um die 15-minütige Lernphase zu starten.

Ihr **Wissens** -Score beträgt 15.00. Der durchschnittliche Wissen-Score in früheren Experimenten war 9.1.	Ihr **Gedächtnis** -Score beträgt 11.00. Der durchschnittliche Gedächtnis-Score in früheren Experimenten war 5.1.

F.6 Description of Test 3

Test 3 looked the same as the first two tests and contained 20 sets of three cities each. Within each set participants had to order cities according to their numbers of inhabitants. A summary of the instructions and the payment scheme was given.