Determinants of the Motivation to Learn: Confidence in Levels and Confidence in Gains*

Mira Fischer and Dirk Sliwka

February 25, 2016

Preliminary draft. Please do not cite or circulate.

Abstract

We analyze the different effects of confidence in prior knowledge and confidence in the ability to learn on the incentives to make human capital investments in a lab experiment. To study the causal effects of the two dimensions of confidence on motivation to learn, we exogenously vary feedback influencing the beliefs in the level of prior knowledge and the ability to learn. We find that confidence in learning ability raises learning investments irrespective of the prior level of knowledge, whereas confidence in prior knowledge has a negative effect on the investments and outcomes of individuals with above average prior knowledge and a positive effect on the investments and outcomes of individuals with below average prior knowledge.

1 Introduction

Motivational beliefs are held to be an important determinant of success in education. Folk wisdom tells us that holding a very favorable opinion of one's abilities may often breed failure as it tempts us to rest on our laurels and lowers our motivation to work hard towards our goals. However, there is a great number of self-help books claiming that increasing our confidence makes us more likely to be successful in life[1] and in educational settings it is often claimed that optimistic beliefs about own ability affect motivation positively. Whereas the economics literature mostly emphasizes the negative effects of too much confidence, many school-related websites and workshop offers claim that fostering children's confidence will improve their motivation to learn. Sometimes teachers and parents are advised to "celebrate a child when it does a great job". However, there seems to be disagreement about whether praise for performance, effort or progress is best to raise confidence and motivation to learn.

*Financial support of the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged. Contact: Mira Fischer: mira.fischer@uni-koeln.de, Dirk Sliwka: dirk.sliwka@uni-koeln.de, University of Cologne

[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.

A straightforward conjecture is that much of the disagreement in the popular discourse about the relationship between feedback, confidence, and performance is caused by the tendency to lump together different types of beliefs into the category of “confidence”. Different types of feedback may influence different confidence beliefs, some of which likely raise motivation to exert effort whereas others likely do not. Indeed, the literature in psychology indicates that there is mixed evidence on the effects of different types of feedback on performance (Kluger and DeNisi, 1998; Hattie and Timperley, 2007). However, little evidence exists on the relationship between feedback, confidence, effort, and performance. Confidence beliefs have featured promimently in the recent literature on economics of education, often subsumed with related concepts under the notion of “soft skills” or “noncogntive skills” that can be shaped by education and are complements to skills measured by standardized tests (Heckman et al., 2006; Heckman and Kautz, 2012; Koch et al., 2015). There is also empirical evidence that less advantaged 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. However, to our knowledge all evidence on the relationship between confidence and performance is correlational and it is not clear to what extent the links are indeed causal.

In the context of education, "confidence" is often used to refer to one’s perceived level of ability (Dweck, 1986; Beckmann et al., 2009) as in many economic studies (Camerer and Lovallo, 1999; Hoelzl and Rustichini, 2005; Stotz and von Nitzsch, 2005; Blavatskyy, 2009) but it is also used to refer to beliefs related to the effectiveness of effort, such as perceived learning ability (Betz and Hackett, 1983; Matsui et al., 1990; Sander and Sanders, 2009). Some types of feedback are more likely to influence the perceived level of ability whereas others are more likely to influence the perceived effectiveness of learning effort and we should not expect that both changes in beliefs will have the same effect on motivation to learn. Economists who have studied confidence in experimental settings have mostly conceived of it to refer to a belief about one’s ability relative to one’s true ability or relative to others’ ability and have focussed on the effects of “overconfidence”, or believing that one’s ability is greater than it actually is. This study, however, focusses on the effects of *raising* or *lowering* confidence on motivation and outcomes. Hence, it does not study the effects of different sizes of confidence *bias* but aims at determining the *marginal* effects of confidence beliefs. This question has been addressed theoretically by Benabou and Tirole (2002) and Benabou and Tirole (2003). We address it experimentally but unlike Benabou and Tirole (2002) and Benabou and Tirole (2003) we will treat confidence as multidimensional, i.e. referring to beliefs in different skills.

We analyze the different effects of *confidence about prior knowledge* and *confidence in the ability to learn* on the incentives to make human capital investments. In order to do so, we invited students to an experiment and allowed them to choose how intensively to prepare for a test. They passed the test and earned a reward if their performance reached a certain threshold. To study the causal effects of the two dimensions of confidence on motivation to learn and test outcomes, we exogenously vary feedback influencing the beliefs in the level of prior knowledge and the ability to learn and then use the exogenous component of the feedback to generate instrumental variables.

The paper makes two contributions, first, to the theory on confidence and, second, to experimental methodology. First, it show 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 understand the effects of confidence on motivation to exert effort, and on learning in particular, we have to understand how effort and the particular ability a confidence belief refers to interact in achieving a goal. An important implication from this is also that school interventions to raise confidence should be carefully designed and evaluated, because they might affect several beliefs that interact differently with motivation to learn. Second, we show how the *causal* effect of beliefs on effort can be studied and thus, for instance, can rule out that psychological dispositions that may be correlated with confidence drive the association between confidence, motivation to exert effort, and performance. Both the use of students as experimental subjects and a realistic high-stakes quiz task rewarded on relative performance let the setting we study, although placed in a computerized laboratory, resemble real exam situations. Our study thus links the economic literature on experiments in education to the literature on motivational beliefs and noncognitive skills.

The remainder of the paper is structured as follows. Section 2 summarizes the related literature on the determinants of effort provison 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 Learning Effort: Incentives, Feedback, and Beliefs About Abilities

Human capital theory has mostly employed the production function paradigm and has studied the relationship between educational inputs, such as class size (Angrist and Lavy, 1999; Hoxby, 2000) and teacher quality (Hanushek and Rivkin, 2010), and eductional and labor market outcomes. However, it has until recently rarely been applied to study the educational process itself, of which student decision making and effort provision, in the presence of incentives and feedback about performance, is a major part (Bishop, 2006). A small but influential strand of literature on education focusses on student behavior from the perpective of microeconomic and behavioral economic theory (Lazear, 2001; Fraja and Landeras, 2006) and its empirical (e.g. Figlio and Lucas 2004; Bonesronning 2008) and experimental tests (e.g. Angrist and Lavy 2009; Fryer 2011; Bettinger 2012; Levitt et al. 2012; Bigoni et al. 2015; Jalava et al. 2015).

So far, most of the field experiments concerned with improving educational test outcomes have focussed on experimental variation of short-term monetary and non-monetary incentives, assuming that suboptimal investments of students in their education is largely driven by hyperbolic discounting of education's greatly delayed rewards (Levitt et al., 2012).

Some field experiments in higher education took a different approach and did not study the effects of incentives but of feedback on student outcomes. For example, field experiments studied the effect of rank-feedback (Tran and Zeckhauser, 2012; Azmat et al., 2015) and natural field experiments studied the effects of relative (Azmat and Iriberri, 2010) and individual grade feedback (Bandiera et al., 2015). The experimental economic literature on tournaments has found both in field (Barankay, 2011; Allcott and Rogers, 2014) and in laboratory studies (Eriksson et al., 2009; Hannan et al., 2008; Ederer, 2010; Guertler and Harbring, 2010; Kuhnen and Tymula, 2012) that the presence of feedback about one's past performance relative to others has a great effect on future decisions in the same setting. However, these experiments in educational settings, work environments, and laboratories focus on comparing feedback with the absence of feedback and none of them show how the effects of feedback on effort are caused by effects of feedback on beliefs about abilities.

People have great uncertainty about their level of ability and productivity, or more generally, the costs and benefits of their actions (Benabou and Tirole, 2002; Falk et al., 2006). This is the reason why feedback about relevant abilities has the power to influence how people react to incentives. There are two major strands of theory dealing with peoples' self-related empircal beliefs that influence expectations of success and failure. The first strand focusses on self-efficacy and locus of control, and the second focusses on overconfidence. We will give a short overview of the cental tenats and findings based on the first and then give a more detailed overview of the literature concerned with the second.

Self-efficacy and locus of control are related psychological constructs important to explaining work motivation (Latham and Pinder, 2005). Having an internal locus of control implies the belief that future outcomes relevant to oneself can be strongly affected by the way one behaves (Eccles and Wigfield, 2002), whereas strong self-efficacy means that one believes that one will be able to behave in the way required to affect future outcomes in the desired way (Bandura, 1977). Both strong self-efficacy and internal locus of control imply that effort spent to attain a goal is perceived as effective in attaining it or making it more likely. Based on U.S. data Heckman et al. (2006), Cebi (2007), and Heckman and Kautz (2012) and based on German data Heineck and Anger (2010) find that internal locus of control predicts labor market outcomes positively. Based on German data Caliendo et al. (2015) find that unemployed individuals with an internal locus of control search for a job more intensively. Stajkovic and Luthans (1998) find a correlation of 0.38 of self-efficacy and work related performance in a meta-analysis of 114 studies.

There are different notions of "confidence" commonly found in the economics literature and they are usually elicited using different methods. Three common methods of measuring confidence (needed in combination with true performance to determine overconfidence) are (1) incentivizing the estimate about how many of a given number of items one solved or will solve correctly, (2) incentivizing the rank belief about how one's past or future performance compares to other's performance, and (3) incentivizing confidence intervals around own estimates. In psychological studies, confidence is sometimes measured using method (1) (Beckmann et al., 2009) but is more often measured by self-assessed questionnaires (Betz and Hackett, 1983; Matsui et al., 1990; Sander and Sanders, 2009). In the latter case the notions of "confidence" and "self-efficacy" or "control beliefs" are often used interchangeably.

Most economic studies on overconfidence have focussed on belief biases to explain outcomes in situations where a choice does not entail a decision about effort (e.g.Odean 1999; Malmendier and Tate 2005). However, some have examined the effect of (over)confidence in situations where effort choices are relevant because outcomes are dependent on effort. In the following we will summarize important findings of research concerned with the effects of (over)confidence on effort. Compte and Postlewaite (2004) study a model in which confidence represents the subjective probability of success at some activity. In their model, overestimating the probability of success has both a negative and a positive effect. The negative effect results from the agent undertaking activities that he should rather not undertake whereas the positive effect results from "positive emotions" while undertaking the activity such that the probability of success rises. Falk et al. (2006) study a laboratory experiment in which subjects are either high types or low types with different probabilities of success when searching. They find that people are substantially uncertain about whether they are a high type or a low typ but as, based on search outcomes, confidence that one is a high type falls, people are less likely to search. Santos-Pinto (2008) theoretically studies the effects of a worker's overestimation of his ability on the the principal's profit from offering pay for performance. He finds that, when effort is not observable, a worker's overestimation of his ability is positive for principals when ability and effort are complements but not when they are substitutes. Herz et al. (2013) study the different effects of overoptimism and judgemental overconfidence on innovative activity. In their framework, overoptimism is related to overestimating the average profit of exploration (versus exploitation) whereas judgemental overconfidence is related to underestimation of the variance of the profits of exploration. They elicit both biases and show both theoretically and experimentally that the first bias leads to too much exploration whereas the second bias leads to too little exploration.

With respect to the model of confidence we propose, the most closely related ideas are found in Spinnewijn (2015) and in Ederer (2010) and Caliendo et al. (2015). Spinnewijn (2015) studies a model in which a two-dimensional choice can be affected by biased beliefs in two dimensions: 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. In his framework, baseline beliefs affect the agent's willingness to save for unemployment whereas control beliefs affect the agent's willingness to exert effort in job search. Ederer (2010) theoretically studies the effect of interim feedback on effort and compares a model where ability plays no role in productivity or enters additively with a model in which effort and ability are complements. Similarly, Caliendo et al. (2015) compare two models of the effects of locus of control on job search. In the first model, individuals with an internal locus of control have a higher subjective probability of receiving a job offer at any given level of search intensity because they believe the payoff to search is higher. In the second model, locus of control is a component of overall perceived ability and individuals think they have a higher baseline probability of being offered a job. Their empirical findings suggest that in their sample the first model is correct but not the second because individuals with a more internal locus of control exert more search effort. Unlike Ederer (2010) and Caliendo et al. (2015), we combine baseline and control belief in one model explaining effort and unlike in Spinnewijn (2015) in our framework both the baseline (prior knowledge) and the control belief (ability to learn) influence a uni-dimensional choice to exert learning effort.

Since we are interested in studying beliefs relevant to how much someone invests in preparing for a high-stakes test that is evaluated relative to others, we elicit confidence incentivizing the rank belief about one's recent performance in a knowledge test closely resembling the high-stakes test in structure and content. That means that we use method (2) described on page 5 to measure the confidence in the level of prior knowledge relative to others. Furthermore, we also use method (2) with respect to a test measuring one's learning ability based on information that is random (such that prior knowledge cannot help) but of the same structure as the learning material for the high-stakes test. So we use method (2) twice to elicit both confidence in prior knowledge and confidence in learning ability relative to others. Both the belief in the relevance of effort (confidence in prior knowledge) and the belief in the effectiveness of effort (confidence in learning ability) are components of the perceived functional relationship between effort and the probability of attaining desired outcomes. Treating the two beliefs separately rather than reducing them to one variable, allows us to study the effects of different types of feedback on behavior. Our design thus allows us to measure incentive compatibly both the perceived level of ability, 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-incentivized questionnaires.

Although incentive compatible measurement of confidence is common in economic laboratory studies, there are very few who have generated *exogenous* variation in beliefs. Furthermore, the causal effect of confidence on *effort* has not been studied yet. To our knowledge, there are two lab experiments that have studied the causal effect of beliefs on behavior and they have done so in settings that do not require effort provision. Mobius et al. (2011) repeatedly give noisy feedback about whether one performed in the better or the worse half in an IQ test to participants, who know that this feedback is correct with a probability of 75%. 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 their ability and find that less confident subjects are causally more likely to be averse. Costa-Gomes et al. (2014) use a trust game and 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 transfer shares.

A real-world example for the situation we study theoretically and experimentally is an agent who faces the decision how much to invest in preparation for a pass-fail test that has high stakes attached for the agent, like admission to a desired educational track or award of a diploma, where the agent is uncertain about how good her knowledge of the subject and her learning ability are. Section 3 will outline our framework and derive optimal responses and hypotheses we test experimentally.

3 The Model

Consider the following simple illustrative model. A risk neutral agent i 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_i and the acquired knowledge Δ_i. Knowledge acquisition is costly and the agent's cost function is

$$\gamma \cdot c\left(\Delta_i\right).$$

The agent is uncertain about both her prior knowledge k_i and the costs to acquire further knowledge γ_i. She knows that both are distributed accoding to the cumulative distribution functions $F_\gamma\left(\gamma_i\right)$ and $F_k\left(k_i\right)$. The agent receives informative signals $s_i = \left(s_{i\gamma}, s_{ik}\right)$ sucht that $\frac{\partial E[\gamma_i|s_{ic},s_{ik}]}{\partial s_{i\gamma}} > 0$ and $\frac{\partial E[k_i|s_{ic},s_{ik}]}{\partial s_{ik}} > 0$. Note that we can decompose

$$\begin{aligned}\gamma_i &= E\left[\gamma_i \left| s_{i\gamma}, s_{ik}\right.\right] + \varepsilon_{\gamma s} \\ k_i &= E\left[k_i \left| s_{i\gamma}, s_{ik}\right.\right] + \varepsilon_{ks}\end{aligned}$$

where $\varepsilon_{\gamma s}$ and ε_{ks} are uncorrelated with the signals $\left(s_{i\gamma}, s_{ik}\right)$ and have mean zero (by the law of iterated expectations).[2] Assume that $\varepsilon_{\gamma s}$ and ε_{ks} have unimodal densities with $g'_{\varepsilon_{\gamma s}}(0) = g'_{\varepsilon_{ks}}(0) = 0$. For ease of notation denote the conditional expectations as

$$\begin{aligned}\hat{k}_i &= E\left[k_i \left| s_{i\gamma}, s_{ik}\right.\right] \\ \hat{\gamma}_i &= E\left[\Delta_i \left| s_{i\gamma}, s_{ik}\right.\right]\end{aligned}$$

such that $\hat{k}_i$ and $\hat{\gamma}_i$ describe the agent's own belief in her knowledge and costs of knowledge acquisition respectively.

[2]To see, for instance, that $Cov\left[s_{i\gamma}, \varepsilon_{\gamma s}\right] = Cov\left[s_{i\gamma}, \gamma_i - E\left[\gamma_i \left| s_{i\gamma}, s_{ik}\right.\right]\right] = 0$ note that $E\left[s_{i\gamma}\left(\gamma_i - E\left[\gamma_i \left| s_{i\gamma}, s_{ik}\right.\right]\right)\right] = E\left[E\left[s_{i\gamma}\left(\gamma_i - E\left[\gamma_i \left| s_{i\gamma}, s_{ik}\right.\right]\right) \left| s_{i\gamma}, s_{ik}\right.\right]\right] = E\left[s_{i\gamma} E\left[\left(\gamma_i - E\left[\gamma_i \left| s_{i\gamma}, s_{ik}\right.\right]\right) \left| s_{i\gamma}, s_{ik}\right.\right]\right] = 0$.

The agent can attain a certain educational outcome, such as an admission to a specific study program, the award of a specific title etc and attains this outcome if $k_i + \Delta_i$ exceeds a threshold value τ. In that case she will receive a reward B. The agent's objective function is thus

$$\max_{\Delta_i} E\left[I_{\{k_i+\Delta_i>\tau\}} B - \gamma c\left(\Delta_i\right) | s_{i\gamma}, s_{ik}\right].$$

Optimal learning efforts:

We can rewrite the agent's objective function as

$$\max_{\Delta_i} E\left[I_{\left\{\hat{k}_i+\varepsilon_{ks}+\Delta_i>\tau\right\}} B\right] - E\left[\left(\hat{\gamma}_i + \varepsilon_{\gamma s}\right) c\left(\Delta_i\right)\right]$$

or

$$\max_{\Delta_i} \Pr\left(\varepsilon_{ks} > \tau - \hat{k}_i - \Delta_i\right) B - \hat{\gamma}_i c\left(\Delta_i\right)$$

which is equivalent to

$$\max_{\Delta_i} \left(1 - G_{\varepsilon_{ks}}\left(\tau - \hat{k}_i - \Delta_i\right)\right) B - \hat{\gamma}_i c\left(\Delta_i\right).$$

We can show:

Proposition 1 *If $\hat{\gamma}_i$ is sufficiently large, there is a unique global optimum $\Delta_i^*\left(\hat{\gamma}_i, k_i\right)$ characterized by*

$$g_{\varepsilon_{ks}}\left(\tau - \hat{k}_i - \Delta_i\right) B - \hat{\gamma}_i c'\left(\Delta_i\right) = 0.$$

(i) Efforts are then strictly increasing in the agent's confidence in her ability to acquire knowledge at low costs $1/\hat{\gamma}_i$.
(ii) Efforts are strictly increasing in the agent's confidence in knowledge $\hat{k}_i$ if and only if $\hat{k}_i$ is smaller than a cut-off value and otherwise strictly decreasing.

Proof:

The first derivative of this function is $g_{\varepsilon_{ks}}\left(\tau - \hat{k}_i - \Delta_i\right) B - \hat{\gamma}_i c'\left(\Delta_i\right)$ and its second derivative is $-g'_{\varepsilon_{ks}}\left(\tau - \hat{k}_i - \Delta_i\right) B - \hat{\gamma}_i c''\left(\Delta_i\right)$. Hence, the function is strictly concave if

$$\frac{\max_\varepsilon \left|g'_{\varepsilon_{ks}}\left(\varepsilon\right)\right|}{\min_\Delta c''\left(\Delta\right)} B < \hat{\gamma}_i.$$

In that case the first order conditions yields the unique optimal effort level.

By implicit differentiation we obtain

$$\frac{\partial \Delta_i^* (\hat{\gamma}_i, k_i)}{\partial \hat{\gamma}_i} = -\frac{-c' (\Delta_i)}{-g'_{\varepsilon_{ks}} \left(\tau - \hat{k}_i - \Delta_i\right) B - \hat{\gamma}_i c'' (\Delta_i)} < 0.$$

And

$$\frac{\partial \Delta_i^* (\hat{\gamma}_i, k_i)}{\partial \hat{k}_i} = -\frac{-g'_{\varepsilon_{ks}} \left(\tau - \hat{k}_i - \Delta_i\right) B}{-g'_{\varepsilon_{ks}} \left(\tau - \hat{k}_i - \Delta_i\right) B - \hat{\gamma}_i c'' (\Delta_i)} \tag{1}$$

Hence,

$$\frac{\partial \Delta_i^* (\hat{\gamma}_i, k_i)}{\partial \hat{k}_i} > 0 \Leftrightarrow g'_{\varepsilon_{ks}} \left(\tau - \hat{k}_i - \Delta_i\right) < 0$$

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

$$\tau > \hat{k}_i + \Delta_i^* (\hat{\gamma}_i, k_i).$$

The right hand side is strictly increasing k_i as $\frac{\partial \Delta_i^*(a_i, k_i)}{\partial k_i} > -1$. To see the latter, note that

$$\frac{\partial \Delta_i^* (\hat{\gamma}_i, k_i)}{\partial k_i} = -\frac{-g' \left(\tau - \hat{k}_i - \Delta_i\right) B}{-g' \left(\tau - \hat{k}_i - \Delta_i\right) B - \hat{\gamma}_i c'' (\Delta_i)} > -1 \Leftrightarrow$$

$$g' \left(\tau - \hat{k}_i - \Delta_i\right) B < g' \left(\tau - \hat{k}_i - \Delta_i\right) B + \hat{\gamma}_i c'' (\Delta_i)$$

which always holds. Hence, condition (1) holds for sufficiently small k and will not hold above a threshold level..

Also:

need additional assumption that guarantees that $\tau > k_i + \Delta_i^* (a_i, k_i)$ for $k = 0$

$$\tau > \Delta_i^* (a_i, 0)$$

recall FOC at $k = 0$

$$g_{\varepsilon_{ks}} (\tau - \Delta_i) B - \hat{\gamma}_i c' (\Delta_i) = 0$$

If the objective function is strictly concave, a sufficient condition for $\tau > \Delta_i^* (a_i, 0)$ is that the objective function is downward sloping at $\Delta_i = \tau$ or

$$g_{\varepsilon_{ks}}(0)\,B - \hat{\gamma}_i c'(\tau) < 0$$

$$g_{\varepsilon_{ks}}(0)\,B < \hat{\gamma}_i c'(\tau)$$

Hence, (1) *a higher confidence in the ability to learn always leads to higher learning investments* as it lowers the perceived marginal costs of learning efforts. However, (2) *confidence in prior knowledge has a positive effect only for agents with a low prior confidence but reduces the incentives to learn for those with a higher prior knowledge.* The intuition is the following: if an agent has a 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, in turn 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.

4 Experimental Method

We invited students to the Cologne Laboratory for Economic Research.[3] Students were not aware of the type of experiment. Upon arrival, registered participants were checked in, randomized to one of two groups and went to their allocated computer. Between computers partition walls prevented participants to look at other participants' screens. Between the two groups, the only difference was the order of the two tests so that we could control for possible ordering effects. (Randomization is explained below.) Before the experiment started, students were informed that they were prohibited 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 whole session. Participants were informed that they would receive the regular showup fee of 2.50 euros and that they could earn additional money during the experiment.[4]

MEASUREMENT OF ABILITIES: After the introduction, participants saw a description of "test 1", the test they were about to take first, which was either a so called "knowledge test"or "memory test"[5]. Likewise, the second test was preceeded by a detailed description. Both tests were incentivized with a piece rate. After each test, participants were asked how many problems they believe to have solved

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

[4] A detailed description of the experiment's timeline, tests, feedback, and belief elicitation can be found in appendices B and C.

[5] Our memory test closely resembles tests used by psycholgists to test working memory capacity (Wilhelm et al., 2013). Working memory capacity of children is a strong predictor of ability to acquire knowledge and new skills, indepedently of IQ (Alloway and Alloway, 2010). See Ackerman et al. (2005) for an overview.

correctly, then they were asked how many sets they believe 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. 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 perform better than the average of participants in the session who did the tests in the same order as they. They were also informed that they could prepare for this third test.

Feedback stage: Participants were explained 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". Then they were explained in easy language how the feedback scores are computed and what can be learned from them. Each score is 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 (cf. Grossman and Owens 2012; Costa-Gomes et al. 2014). The randomly distributed noise term thus creates exogenous variation in feedback about knowledge and learning ability without employing deception. Then the personal feedback scores and average feedback scores of participants in past sessions are displayed on the same screen, which turns the scores into a noisy but informative signal about individual performance relative to others' average performance. [6]

measurement of beliefs: Participants are asked to estimate their rank in the knowledge and in the memory test relative to the half of participants in the room who worked on the two tests in the same order as they. They are informed that they can earn one euro, respectively, for estimating their rank correctly. In order to obtain a precise measure of confidence that is not biased by risk or ambiguity aversion, we decided to pay subjects for exactly right estimates of their rank only. We opted against using the quadratic scoring rule (Selten, 1998), although it is a common elicitation method in economic experiments, because unlike our simple method it is not robust to risk aversion when the range of beliefs is limited such that reporting beliefs towards the ends of the range is riskier than reporting beliefs closer towards the middle. Furthermore, our simple method can be easily explained to subjects. Since the range of beliefs in our context is small due to a limited number of ranks, the chances of exactly right estimates are reasonable and rewarding those induces a strong incentive to report one's belief accurately.[7]

Investment stage: After participants received their knowledge score and their memory score they are shown a screen where "test 3: combined knowledge and memory test" and how one can prepare for it is described in detail. Participants are explained that this test is based on the same field of knowledge

[6]We opted to always use 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 comparable.

[7]There is controversy about how to properly measure confidence in decision making. Whereas asking people for their beliefs and not giving them an incentive to lie does not induce biases but imprecise reporting of beliefs, rewarding people for their estimates will produce more precise estimates but these can easily be biased by risk aversion (Hoelzl and Rustichini, 2005) or ambiguity aversion (Blavatskyy, 2009). Our simple incentivized method neither suffers from risk nor from ambiguity aversion.

and has the same length and structure as the knowledge test. Furthermore, they are told that they can prepare for it by acquiring test relevant information (a sample piece of information is shown). In order to do so, they would receive a budget of 3 euros independently of their previous performance with which they could buy information in increments of 10 items (or 0.5 euros) which then would be displayed in a 15 minutes learning phase before test 3 starts.

Measurement of outcomes: Participants take test 3. After the test participants fill in a questionnaire. In the very end they are informed how much money they earned (and how they performed) in each stage of the experiment.

5 Experimental Results

Our main interest is in the size of investment that participants make to prepare for test 3 and how this investment is causally related to beliefs about relative learning ability and relative prior knowledge. We begin by studying the correlation between beliefs and investment:

$$Investment_i = \alpha + \beta ConfidenceLearning_i + \gamma ConfidenceKnowledge_i + \zeta_i \quad (2)$$

Figure1 shows plots of quadratic predictions of investment behavior as a function of the respective belief measured in percentile ranks. To facilitate interpretation of coefficients and account for the varying group sizes due to some no-shows, we inverted rank beliefs and standardized them to percentile ranks such that the maximum possible level of confidence is 100 and the minimum possible confidence is 0. As can be seen in the left graph of Figure 1, there is a linear and positive relationship between how good a person thinks her learning ability is and how much she invests in learning. The better she thinks her memory is compared to other people, the larger the amount of costly information she acquires for the study period. The right graph of Figure 1 shows that the relationship between the belief in level of prior knowledge and the investment in studying is hump shaped. Investment is highest if the person thinks that her knowledge is about average.

In the following we will investigate whether the correlation between beliefs and investment is due to a causal effect of beliefs on investment. 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 for beliefs in regressions explaining behavior and outcomes. By doing so, we will only use the exogenous component of beliefs, uncorrelated with individual traits, to explain behavior.

5.1 Causal Effect of Feedback Manipulation on Beliefs

In order to identify the effect of our feedback manipulations on participants' beliefs we first regress our incentivized measures of confidence in learning ability and confidence in knowledge elicited after the feedback on the exogenously varied noise term added to the memory feedback and the noise term added to the knowledge feedback. We also include an interaction term between the two noise terms to capture how the presence of one feedback bias affects the effectiveness of the other feedback bias. We estimate the following specification by OLS:

$$\begin{aligned} Confidence_i = \alpha + \beta NoiseTermMemory_i + \gamma NoiseTermKnowledge_i \\ + \delta NoiseInteraction + \varepsilon Controls + \zeta_i \end{aligned} \quad (3)$$

We control for ability by including the test results and school GPA. Additionally we control for subject of study, gender, number of semesters a student has completed at university and session fixed effects. Since the noise terms were randomly assigned, the inclusion of controls does not bias the coefficients of the noise terms but helps us to gain statistical power in our instrumental variable regressions. For reasons of consistency, we include them in model 2, which is thus equal the first stage of our IV regressions. The results in Table 1 show that our exogenous variation of beliefs indeed worked: the respective noise term has a strong effect on the participants' beliefs about both their memory and their knowledge. A one unit increase in the noise term in the memory feedback on average causes participants to believe that their memory is 7.8 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 6.0 percentile ranks better. Hence, our manipulation worked and the exogenous variation in feedback scores indeed affected beliefs. In the following two subsections, we will study the causal effect of confidence in learning and prior knowledge on investment behavior and test outcomes.

5.2 Causal Effect of Beliefs on Learning Investments

By studying whether our treatment affected behavior through affecting beliefs we can address the question whether the relationships presented in Figure 1 indeed reflect causal effects. This will also allow us to test the hypotheses stated in section 3. In order to do so we run an instrumental variables regression of beliefs on investment where the two beliefs are instrumented by the two noise terms and the interaction between them. The first stage of the IV regression is given by equation 3, the second stage is given by:

$$Investment_i = \alpha + \beta ConfidenceLearning_i + \gamma ConfidenceKnowledge_i + \varepsilon Controls + \zeta_i \quad (4)$$

Column (1) of Table 2 shows that confidence in gains significantly increases investment whereas the effect of confidence in levels of prior knowledge is insignificant when looking at the whole sample. This is what we expected under our hypothesis. Given that the true relationship between confidence in prior knowledge and investment is hump shaped a linear estimator represents a miss-specification. Since we expected a positive effect for individuals with below average prior knowledge and a negative effect for individuals with above average prior knowledge, we split the sample at average outcome of the knowledge test and estimate effects for the worse half and the better half separately. In Columns (2) and (3) we can see that both in the better and in the worse half of participants confidence in learning gains 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 gains we find that an increase of confidence by 10 percentile ranks raises investment in learning by about 7 euro cents for the better half of students and about 9 euro cents for the worse half of students. For confidence in levels we find that an increase of confidence by 10 percentile ranks lowers investment in learning by about 10 euro cents for students with above average level of prior knowledge but raises investment in learning by about 11 Euro cents for students with below average level of prior knowledge.

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 is. We also find strong evidence in favor of the hypothesis that increasing the confidence in prior knowledge reduces incentives for individuals whose knowledge is already above average. Within the group of people whose knowledge is below average, we find evidence that confidence in knowledge increases learning investments as expected.

5.3 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. Of course, looking at the very short run effects of a confidence intervention in the lab should be be done with great caution as it heavily depends on the effectiveness of the learning technology we provided to students. Furthermore, zero effects of a confidence intervention on outcomes in a test test taken 15 minutes after the intervention would not imply that a behavioral change should not affect outcomes in the medium or long term. We use a probit estimation method for endogenous explanatory variables based on Newey (1987) to test whether beliefs causally affect the probability of passing the test. The first stage uses a linear link function and is equal to equation 3. The second stage uses a probit link function and is

$$Pr(y=1|x) = G(\alpha + \beta ConfidenceLearning_i + \gamma ConfidenceKnowledge_i + \varepsilon Controls + \zeta_i) \qquad (5)$$

where

$$G(\mathbf{x}\beta) = \Phi(\mathbf{x}\beta) \equiv \int_{-\infty}^{\mathbf{x}\beta} \phi(v)dv.$$

As can be seen by looking at Table 3, we do find a significant effect of confidence in levels of prior knowledge on outcomes but the effect of confidence in gains is not significant. The results for confidence in gains point in the same direction as the coefficients in Table 3, where we looked at the behavioral effects of the two dimensions of confidence. We conclude from this evidence that as students who have above average levels of prior knowledge become more confident about their level, they invest less effort in learning and become less likely to pass the test. For students with below average levels of prior knowledge we observe the opposite effect: as they become more confident about their level they invest more effort in learning and become more likely to pass the test.

6 Conclusion

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 subjectively moves them closer to the passing threshold and 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 because the perceived marginal productivity of effort increases. Furthermore, we expected increases and decreases in effort due to a change in beliefs to affect test outcomes.

Our results largely support our 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 raising confidence in prior knowledge also show through in test outcomes. Raising confidence in prior knowledge decreases the probability that an individual with above average prior knowledge passes the test, whereas it increases the passing probability of individuals with below average prior knowledge. Our results also indicate that confidence in learning ability has a positive effect on the probability to pass the third test. However, given our lack of statistical power, the estimated effects in this latter dimension are marginally too small to be significant at conventional levels.

We find strong evidence 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. However, the results already show that generalized statements about the role of confidence can be misleading and confidence should be viewed as a multidimensional concept. This leads to insights not only for the design of grading systems in schools and the education of children but also for the design of feedback policies in firms. While raising confidence in the ability to acquire a certain skill can be beneficial, raising confidence in the skill itself can be detrimental. Thus praising people for something they achieved in the past may have drawbacks but praising people for their progress in achievements may be very beneficial as it can motivate them to achieve more in the future.

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.

Allcott, H. and Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104(10):3003–37.

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.

Angrist, J. and Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *American Economic Review*, 99(4):1384–1414.

Angrist, J. D. and Lavy, V. (1999). Using maimonides’ rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 114(2):533 – 575.

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.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2):191 – 215.

Barankay, I. (2011). Rankings and social tournaments: Evidence from a crowd-sourcing experiment. *unpublished working paper*.

Beckmann, N., Beckmann, J. F., and Elliott, J. G. (2009). Self-confidence and performance goal orientation interactively predict performance in a reasoning test with accuracy feedback. *Learning and Individual Differences*, 19(2):277 – 282.

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.

Bettinger, E. P. (2012). Paying to learn: The effect of financial incentives on elementary school test scores. *Review of Economics and Statistics*, 94(3):686 – 698.

Betz, N. E. and Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior*, 23(3):329 – 345.

Bigoni, M., Fort, M., Nardotto, M., and Reggiani, T. G. (2015). Cooperation or competition? a field experiment on non-monetary learning incentives. *The B.E. Journal of Economic Analysis & Policy*, 15(4):1753 – 1792.

Bishop, J. (2006). Drinking from the fountain of knowledge: student incentive to study and learn - externalities, information problems and peer pressure. In *Handbook of the economics of education ; Vol. 2.*, pages 909–944. North-Holland, Amsterdam.

Blavatskyy, P. (2009). Betting on own knowledge: Experimental test of overconfidence. *Journal of Risk and Uncertainty*, 38(1):39–49.

Bonesronning, H. (2008). The effect of grading practices on gender differences in academic performance. *Bulletin of Economic Research*, 60(3):245–264.

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.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1):306–318.

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.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41(10):1040 – 1048.

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.

Eriksson, T., Poulsen, A., and Villeval, M. C. (2009). Feedback and incentives: Experimental evidence. *Labour Economics*, 16(6):679 – 688.

Falk, A., Huffman, D., and Sunde, U. (2006). Self-confidence and search. *IZA Discussion Papers*, 2525.

Figlio, D. N. and Lucas, M. E. (2004). Do high grading standards affect student performance? *Journal of Public Economics*, 88(9-10):1815–1834.

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.

Fraja, G. D. and Landeras, P. (2006). Could do better: The effectiveness of incentives and competition in schools. *Journal of Public Economics*, 90(1 - 2):189 – 213.

Fryer, R. G. (2011). Financial incentives and student achievement: Evidence from randomized trials. *The Quarterly Journal of Economics*, 126(4):1755–1798.

Greiner, B. (2004). *An Online Recruitment System for Economic Experiments*, chapter 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.

Guertler, O. and Harbring, C. (2010). Feedback in tournaments under commitment problems: Experimental evidence. *Journal of Economics & Management Strategy*, 19(3):771–810.

Hannan, R. L., Krishnan, R., and Newman, A. H. (2008). The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. *The Accounting Review*, 83(4):893–913.

Hanushek, E. A. and Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100(2):267–71.

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):pp. 411–482.

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

Herz, H., Schunk, D., and Zehnder, C. (2013). How do judgmental overconfidence and overoptimism shape innovative activity? *CESifo Working Paper Series*, 4084.

Hoelzl, E. and Rustichini, A. (2005). Overconfident: Do you put your money on it? *The Economic Journal*, 115(503):305–318.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics*, 115(4):1239 – 1285.

Jalava, N., Joensen, J. S., and Pellas, E. (2015). Grades and rank: Impacts of non-financial incentives on test performance. *Journal of Economic Behavior & Organization*, 115:161 – 196.

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

Latham, G. P. and Pinder, C. C. (2005). Work motivation theory and research at the dawn of the twenty-first century. *Annual Review of Psychology*, 56:485–516.

Lazear, E. P. (2001). Educational production. *Quarterly Journal of Economics*, 116(3):777 – 803.

Levitt, S. D., List, J. A., Neckermann, S., and Sadoff, S. (2012). The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. 18165.

Malmendier, U. and Tate, G. (2005). Ceo overconfidence and corporate investment. *Journal of Finance*, 60(6):2661–2700.

Matsui, T., Matsui, K., and Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior*, 37(2):225 – 238.

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

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

Odean, T. (1999). Do investors trade too much? *American Economic Review*, 89(5):1279–1298.

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.

Sander, P. and Sanders, L. (2009). Measuring academic behavioural confidence: the abc scale revisited. *Studies in Higher Education*, 34(1):19 – 35.

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

Selten, R. (1998). Axiomatic characterization of the quadratic scoring rule. *Experimental Economics*, 1(1):43–61.

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

Stajkovic, A. D. and Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, 124(2):240 – 261.

Stotz, O. and von Nitzsch, R. (2005). The perception of control and the level of overconfidence: Evidence from analyst earnings estimates and price targets. *Journal of Behavioral Finance*, 6(3):121–128.

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).

Appendix

A Figures and tables

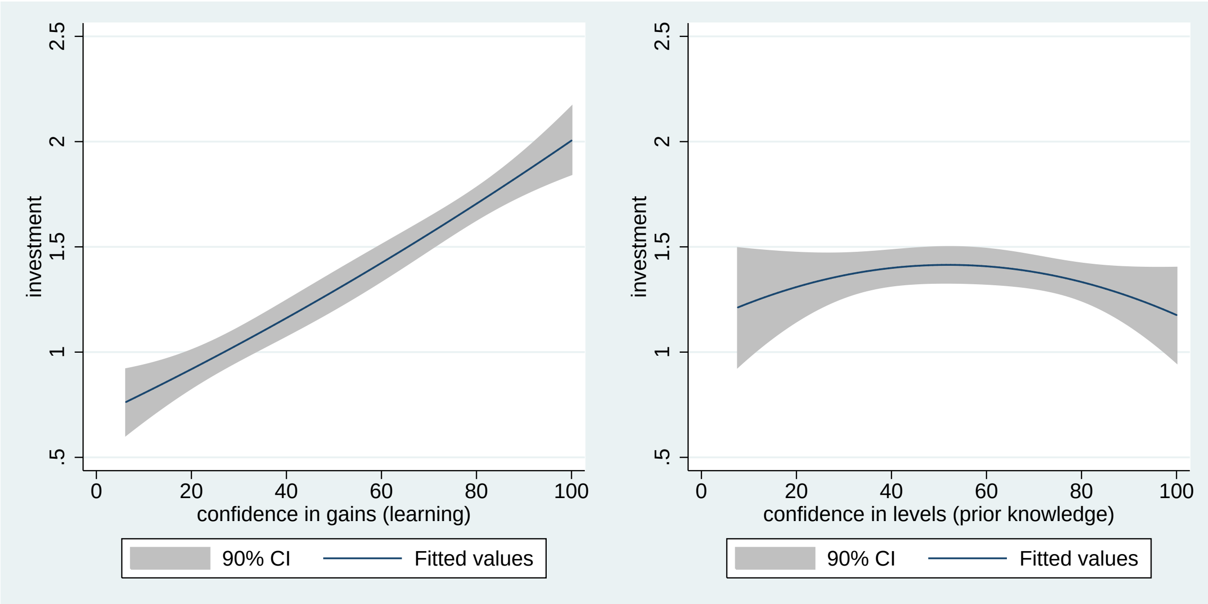


Figure 1: Association of confidence in levels and confidence in gains with investment in learning

Table 1: Table: First Stages

	(1) Belief Memory	(2) Belief Knowledge
Noise Term Memory	7.789***	-0.283
	(17.43)	(-0.66)
Noise Term Knowledge	-0.529	6.020***
	(-1.17)	(13.85)
Noise Interaction	-0.166	0.0676
	(-0.53)	(0.22)
Memory Test	8.526***	-0.299
	(34.66)	(-1.26)
Knowledge Test	-0.415	5.852***
	(-1.43)	(20.89)
Constant	18.56***	13.97**
	(3.02)	(2.36)
R^2	0.752	0.600
Sample Size	615	615

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Confidence on Investment (IV)

	(1) Invest. (All)	(2) Invest. (Better)	(3) Invest. (Worse)
Belief Memory	0.00856***	0.00745*	0.00921**
	(2.68)	(1.69)	(2.27)
Belief Knowledge	-0.00183	-0.00962*	0.0114*
	(-0.42)	(-1.86)	(1.79)
Memory Test	0.0546*	0.0530	0.0620
	(1.84)	(1.27)	(1.63)
Knowledge Test	-0.0183	-0.0665	-0.0759*
	(-0.60)	(-1.52)	(-1.72)
Constant	0.904**	1.641***	1.140*
	(2.55)	(3.41)	(1.80)
R^2	0.270	0.290	0.358
Sample Size	615	353	262
F-Test (weak ID), Memory	106.2	56.67	48.81
F-Test (weak ID), Knowledge	66.37	49.05	30.22

Two-stage least squares estimates; t statistics in parentheses; additional control variables: school GPA, dummy variables for gender, field of study,semester of study, and session; Model 1: whole sample, Model 2: average and above performance in knowledge test, Model 3: below average performance in knowledge test

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

Table 3: Confidence on Prob. of Passing Test 3 (IV)

	(1) Prob. Passing (All)	(2) Prob. Passing (Better)	(3) Prob. Passing (Worse)
Belief Memory	0.00243	-0.00426	0.0110
	(0.48)	(-0.58)	(1.26)
Belief Knowledge	-0.00465	-0.0173**	0.0217*
	(-0.70)	(-2.02)	(1.87)
Memory Test	0.0756	0.143**	0.0194
	(1.57)	(1.97)	(0.24)
Knowledge Test	0.0569	0.0240	-0.0915
	(1.23)	(0.32)	(-1.08)
Constant	0.154	1.952**	-1.126
	(0.28)	(2.04)	(-1.15)
χ^2	59.92	46.51	35.67
Sample Size	608	347	251

IVProbit estimates (Newey's (1987) minimum-chi-squared estimator,"ivprobit, twostep" in STATA); t statistics in parentheses; additional control variables: school GPA, dummy variables for gender, field of study, semester of study, and session; Model 1: whole sample, Model 2: average and above performance in knowledge test, Model 3: below average performance in knowledge test

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

Table 4: Reduced Form (Investment)

	(1) Invest. (All)	(2) Invest. (Better)	(3) Invest. (Worse)
Noise Term Memory	0.0669***	0.0569*	0.0744*
	(2.62)	(1.71)	(1.75)
Noise Term Knowledge	-0.0160	-0.0652*	0.0622
	(-0.62)	(-1.91)	(1.45)
Noise Interaction	-0.00885	-0.0109	-0.0322
	(-0.49)	(-0.47)	(-1.06)
Memory Test	0.128***	0.117***	0.135***
	(9.09)	(6.13)	(6.08)
Knowledge Test	-0.0326*	-0.126***	-0.0351
	(-1.95)	(-3.45)	(-0.86)
Constant	1.032***	1.620***	1.494***
	(2.94)	(2.93)	(2.63)
R^2	0.254	0.325	0.374
Sample Size	615	353	262

OLS estimates; t statistics in parentheses; additional control variables: school GPA, dummy variables for gender, field of study, semester of study, and session; Model 1: whole sample, Model 2: average and above performance in knowledge test, Model 3: below average performance in knowledge test

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

Table 5: Reduced Form (Prob. of Passing Test 3)

	(1) Prob. Passing (All)	(2) Prob. Passing (Better)	(3) Prob. Passing (Worse)
Noise Term Memory	0.0203	-0.0266	0.0775
	(0.52)	(-0.50)	(1.11)
Noise Term Knowledge	-0.0298	-0.109**	0.129*
	(-0.76)	(-2.03)	(1.85)
Noise Interaction	-0.00881	-0.0179	-0.0206
	(-0.32)	(-0.48)	(-0.42)
Memory Test	0.0973***	0.107***	0.103***
	(4.49)	(3.52)	(2.80)
Knowledge Test	0.0285	-0.0711	-0.0111
	(1.12)	(-1.21)	(-0.17)
Constant	0.129	1.511*	-0.559
	(0.24)	(1.67)	(-0.62)
χ^2	65.30	53.78	42.05
Sample Size	608	347	251

Probit estimates; t statistics in parentheses;additional control variables: school GPA, dummy variables for gender, field of study, semester of study, and session; Model 1: whole sample, Model 2: average and above performance in knowledge test, Model 3: below average performance in knowledge test

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

Table 6: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Noise Term Memory	-0.04	1.42	-2	2	615
Noise Term Knowledge	0.01	1.41	-2	2	615
Belief Memory	53.81	28.54	6.25	100	615
Belief Knowledge	58.11	21.59	7.69	100	615
Memory Test	5.13	2.55	0	11	615
Knowledge Test	8.85	2.21	0	16	615
Investment	1.36	0.94	0	3	615
Prob. of Passing Test 3	0.5	0.5	0	1	615
Profit	11.32	5.03	3.2	19.4	615
Female	0.62	0.49	0	1	615
Semester	6.55	3.9	1	23	615
School GPA	2.05	0.6	1	3.5	615
Session	12.82	6.9	1	24	615
Humanities	0.16	0.36	0	1	615
Social Sciences	0.1	0.3	0	1	615
Law	0.05	0.22	0	1	615
Business Administration	0.27	0.44	0	1	615
Economics	0.14	0.35	0	1	615
Medicine	0.05	0.22	0	1	615
Natural Sciences	0.08	0.27	0	1	615
Psychology	0.01	0.12	0	1	615
Other	0.15	0.36	0	1	615

B Timeline of the experiment

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

   - „knowledge test“: participants have to rank sets of three German cities according to their population
   - „memory test“: participants are presented a list of cities with (arbitrary) „cities codes“ which they can memorize, then they have to rank sets of three cities according to these codes.
   - Immediately after each test participants estimate their performance and other‘s average performance in each test (not incentivized)

2. Info on further course: 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 how feedback is computed.

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

4. Belief elicitation: Subjects estimate their rank position in both tests

5. Decision stage: 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: Subjects take the third test (economic outcome variable)

C Details on the tests, feedback, elicitation of beliefs, and decision stage

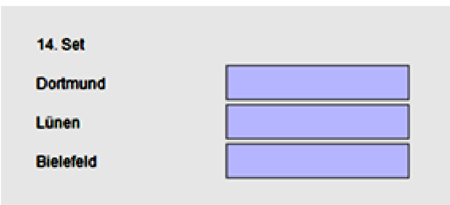
The experiment was conducted in German, so in the following we give the English translation of the texts. All the cities used in the experiment come from a sample of the 200 largest cities in Germany. We pretested all instructions and tests to ensure that they are understandable and produced a sufficient 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.

C.1 Description of knowledge test

The instruction on the introductory screen said:

> "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. 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 also shown. For example, a set of three cities looked like this:



C.2 Description of memory test

The instruction on the introductory sceen 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, respectively, 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 also 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

C.3 Description of feedback

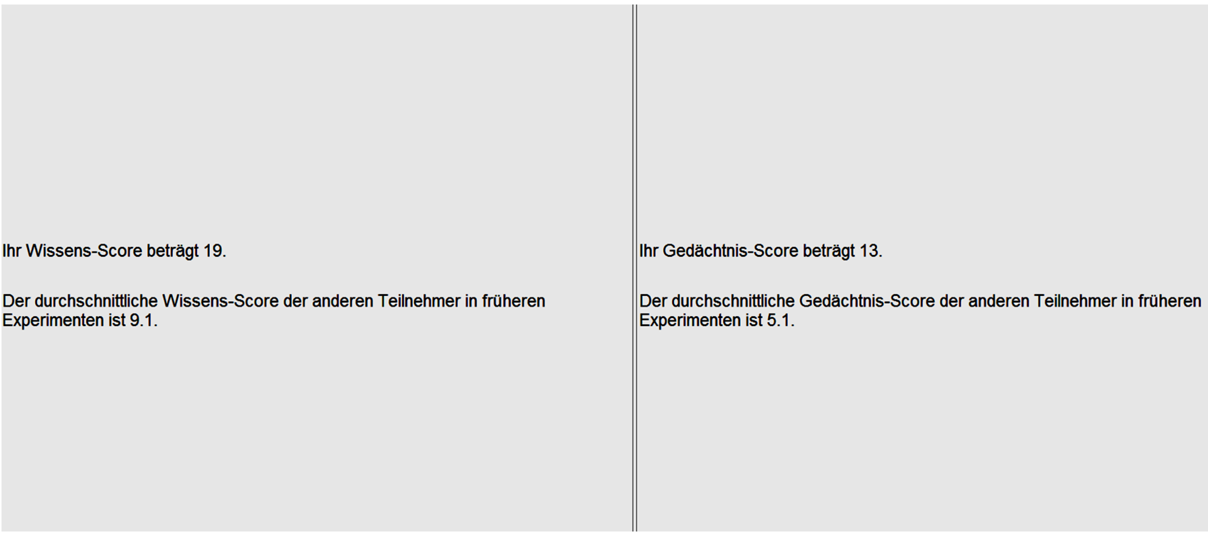
After subjects are told that there will be a third "main test" and that they can prepare for it, they are told that they are about to receive feedback. Next, they are shown a screen where the computation the "feedback scores" is explained:

> "The experimental software will no generate a feedback score and a memory score for earch 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 indepent 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 im each test with a certain imprecision but in expectation equal the actual number of your correct answers."

The feedback screen contained both a participant's two scores and the respective average score of participants in earlier experimental sessions. It looked like this:



C.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 estimate 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 believed ranks in the knowledge and the memory test by clicking on radio buttons.

C.5 Decision stage

The decision screen containd 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 number 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 cities in the knowledge test about the number 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 to possiblity to improve you knowledge of cities in a learning phase.

Description of preparation for test 3

In order to prepare for test 3, you may buy information about cities' number 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 packagages 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 are 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 learn 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.

Ihr **Wissens** -Score beträgt 19.00. Der durchschnittliche Wissen-Score in früheren Experimenten war 9.1.	Ihr **Gedächtnis** -Score beträgt 13.00. Der durchschnittliche Gedächtnis-Score in früheren Experimenten war 5.1.

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.

C.6 Description of test 3

Test 3 looked the same as the first two tests. A summary of the instructions and the payment scheme was given.