Ruhr Economic Papers #328

Christoph Helbach and Klemens Keldenich

Teaching in the Lab: Financial Incentives in the Education Process

Ruhr Graduate School in Economics

Universität Duisburg Essen

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788377
ISSN 1864-4872 (online)
ISBN 978-3-86788-377-1

Christoph Helbach and Klemens Keldenich[1]

Teaching in the Lab: Financial Incentives in the Education Process

Abstract

This study uses a laboratory experiment to analyze the effectiveness of performance-based monetary incentives in the teaching process. The process of knowledge transmission is recreated using a video-stream. Four different teacher payment schemes are compared, three of which depend on the student's success. Furthermore, the experiment is done with two different subject pools: prospective teachers and regular students. Results indicate that prospective teachers do not react to monetary incentives in the given task. However, regular students do react in the expected way: Teachers transmit a significantly higher share of their knowledge when paid according to student performance.

JEL Classification: C91, D03, I21, J33

Keywords: Education; monetary incentives; video analysis

April 2012

1 Christoph Helbach, University of Duisburg-Essen; Klemens Keldenich, University of Duisburg-Essen and Ruhr Graduate School in Economics. *– We thank Jeannette Brosig-Koch as well as participants at the ESA World Meeting 2011, the ESA European Meeting 2011, the GfeW Annual Conference 2011, the Nordic Conference on Behavioral and Experimental Economics 2011 and seminar participants at the University of Duisburg-Essen for helpful comments and suggestions. Financial support by the Fritz Thyssen Stiftung is gratefully acknowledged. Beatus Ille, Wolfgang Lambertz, and Thomas Lanko provided excellent research assistance.* *– All correspondence to: Klemens Keldenich, RGS Econ, c/o Chair for Economic Policy: University of Duisburg-Essen, Faculty of Economics, Universitätsstr. 12, 45117 Essen, Germany, Email: klemens.keldenich@ibes.uni-due.de.*

1 Introduction

The transfer of knowledge is an important part of many economic and everyday situations. Examples can be easily found in the corporate sector (when new employees need to receive firm-specific information from their colleagues), in private life (when parents give advice to their children career choice, for example), and most prominently in the education sector (which is in essence mainly occupied with the transfer of knowledge). While standard economic theory maintains that people need incentives to exert effort, little is known about the interplay of incentives and effort in knowledge transmission processes. In particular the role of monetary incentives in this context is far from clear even though there can be little doubt that financial considerations play a big role at least in professional environments. The paper presented here aims to shed some light on this aspect by analyzing the role of monetary incentives in knowledge transfer, using a laboratory experiment.

Although in general both agents in a knowledge transmission process - the provider and the receiver of knowledge - can be subject to incentive problems, it is particularly important to incentivize knowledge providers as they usually do not have an inherent incentive to spend effort on the task. New employees for example may be motivated to learn with a view to increasing their future opportunities. Their advisors on the other hand may shirk, for instance because they concentrate their effort on incentivized tasks to increase their monetary payoff. In such a scenario the knowledge provider should receive a reward for her efforts that should ideally be dependent on the amount of knowledge the intended recipient actually receives. This incentive-compatible approach is however rarely used. School teachers for example often receive with a fixed wage, regardless of their students' performance[1]. Similarly, companies seldom include knowledge transfer in the criteria used for variable compensation.

In many industries, monetary incentive schemes are however a standard way to motivate employees in areas apart from knowledge transfer. Adams et al. (2009) estimate that in 2002, 33% of the Fortune 1000 companies used some kind of individual incentive. The economic argument for variable pay linked to performance is to align the interests of employers and employees (Kessler and Purcell, 1992). Existing research has shown that monetary incentives indeed affect agents' effort in many settings. Several studies have found a positive effect of performance pay on worker effort and productivity. This effect is however task dependent and Camerer and Hogarth (1999) conclude in a survey of experimental results that "The data show that incentives sometimes improve performance, but often don't" (pg. 32). In another survey Prendergast (1999) looks at the provision of monetary incentives in firms and concludes that "incentives matter" (pg. 11).

There is, however, also evidence that the introduction of monetary incentives might have adverse

[1] This is e.g. true for most teachers Germany, who do not receive any variable wage at all.

effects: Gneezy and Rustichini (2000) show that an extrinsic source of motivation like money might crowd out intrinsic motivation to do the task well and therefore diminish effort. As a consequence low monetary incentives reduce performance in comparison to a scenario without monetary incentives. In contrast, Pokorny (2008) finds that even low piece rate payments improve performance in a real effort experiment. Despite these diverging key results both studies provide evidence for non-monotonic effects of monetary incentives. They thus exemplify that finding the optimal incentive scheme is not a trivial exercise. This is even more so if, apart from possible crowding out effects, performance is hard to measure, as it is the case with knowledge transfer or learning success. Adams et al. (2009) give several examples where this problem leads to a worse outcome compared to a situation without conditional monetary incentives and Prendergast emphasizes that "there has been insufficient focus on workers whose outputs are hard to observe" (pg. 11).

To the best of our knowledge, this paper is the first to analyze the actual teaching process in the laboratory. There is, however, a literature on information transmission by way of advice giving (e.g. Chaudhuri et al., 2006; Schotter and Sopher, 2007). This literature has shown that participants in experiments strongly respond to advice, even in an anonymous experimental setup. These studies differ however from the paper at hand in at least two significant ways. First, advice deals with situations such as public good or ultimatum games. Thus, individual preferences and interaction play a role and there is no such thing as correct advice. Second, the focus is the advice giving environment and incentive schemes are not alternated as treatment variables.

While there is no laboratory evidence on monetary incentives in knowledge transmission processes, recent studies from the field of education economics have provided field evidence on teacher performance pay. The most compelling evidence stems from field experiments in India (Muralidharan and Sundararaman, 2011) and Israel (Lavy, 2002, 2009). The data from these experiments reveal a significantly positive effect of merit pay systems. Figlio and Kenny (2007) as well as Woessmann (2011) draw a similar conclusion from analyzing survey data from the United States and cross-country data, respectively. Despite those positive results there is also evidence that sheds some doubt on teacher performance pay. Reback (2008), Eberts et al. (2002), and Martins (2009) find some negative effects of teacher performance pay by analyzing data from the United States, and Portugal, respectively. The study by Reback in particular provides useful insights. He analyzes the effect of the American "No Child Left Behind Act" of 2001 which penalizes schools which do not meet minimum proficiency requirements. The author finds that teachers as a consequence focus on those students close to this minimum proficiency threshold which results in an improvement only for relatively low performing students. Simlarly, Glewwe et al. (2010) find that teachers focus on incentivized goals and disregard non-incentivized ones.

The brief literature overview shows that a final understanding of the effects of monetary incentives

requires further research. The present paper aims to provide one further step in this direction by analyzing instructors' incentives in a knowledge transmission laboratory experiment. It contributes to the literature in various ways. First, the controlled laboratory environment reduces problems of measuring effort by providing observable outcomes. Second, the experimental setup allows a direct comparison of four different incentive schemes at reasonable costs. Third, prospective teachers can be compared to other participants[2] to draw preliminary conclusions on whether they respond differently to monetary incentives in this specific task.

2 Experimental design

2.1 Structure

All subjects in the experiment are assigned to one of two roles, instructor or pupil, which they keep throughout the whole experiment. The experiment consists of three parts. All subjects are placed in sound protected cabins with closed doors during the experiment. Subjects are allowed to make notes but are asked to use the notepad provided by the experimenters. The instructions include information about all three parts of the experiment[3].

In the first part, only subjects in the instructor role (from now on "instructors") participate. After reading the instructions, they see a presentation on two topics. The instructors can freely move through the slides of this presentation, but only have 20 minutes time to do so. The two topics used in the presentation are a self-developed card game called "Pizzabäcker"[4] and the artificial language "Lojban"[5]. These two topics have been chosen because they fulfill several requirements: No subject should have prior knowledge of them, the topics should yield enough questions to allow variance, they should be easy to translate into multiple choice questions, and sufficiently complex to explain and teach. Two topics rather than only one were chosen to decrease the influence of idiosyncratic capabilities of the participants (if, e.g., one subject is very talented at understanding languages, the addition of a second unrelated topic decreases the distorting influence of this). Both topics take up the same number of slides in the presentation, which is known by the participants beforehand. In addition to the presentation itself, instructors also receive additional material to help them understand the topics better: a complete set of cards for the card game and a sheet containing the alphabet and one sentence for the language. During the presentation for the instructors, the subjects in the pupil role (from now on "pupils") also enter the laboratory. Every pupil is randomly assigned to one instructor. The pupils are given the same instructions as the instructors

[2] See Section 2.2 for a detailed explanation.
[3] A set of complete instructions can be found in the Appendix.
[4] The card game's name translates to "Pizza baker" and was adapted from the game "Dia de los muertos". The original game was created by Frank Graham, who graciously allowed us to use it.
[5] See http://www.lojban.org for a short introduction.

(including the material belonging to the card game and the language). While the pupils read those instructions, the instructors finish the presentation and then have 10 minutes time to prepare for the second part of the experiment.

In the second part, the actual teaching takes place. For each instructor-pupil pair a one-way video conference is established, such that the pupil can see and hear the instructor but not vice versa. This one-sided mode of communication was chosen to rule out the possible influence of the pupil on the quality of the teaching. If a pupil is very smart or motivated, she can positively influence the teaching quality and effort by, e.g., asking good questions. As the aim of the design is to isolate the effects the treatments have on the instructors, this is not desirable. The connection between pupil and instructor lasts for 10 minutes, and the resulting video is recorded with the knowledge (and prior consent) of the subjects. Before the 10 minutes start, instructors are given some extra time to calibrate the position of the camera, their headset and the volume. During the 10 minutes, pupils can only press a button which indicates to the instructor that they can see and hear her properly. Instructors are not given any encouragement on how to spend the 10 minutes, they can e.g. easily avoid teaching altogether and concentrate on their notes instead (some instructors do in fact point the camera away from them so their pupil cannot see them).

In the third and final part, instructors and pupils all have to answer the same 30 multiple choice questions (15 for each topic), where four possible answers are given for each question and exactly one answer is correct. Each question is displayed for 40 seconds on the screen, the subjects cannot speed up the questions or return to older questions. This procedure is described in the instructions for everybody. After the questions are answered, every subject is informed how many questions and which questions he has answered correctly. In the treatments where instructors' payoff depends on the pupils' answers, the instructors are additionally informed about the number of correct answers of their assigned pupil.

After the experiment itself, subjects fill out a questionnaire with questions about demographics (sex, age, study length, study subject), their school grades (last math grade and last German grade), some personality measures, teaching experience, general card game experience, risk attitude, and if they knew "Lojban" or the card game before the experiment.

2.2 Treatments

Pupils always receive the same payoff: For each correctly answered question, they receive EUR 0.75. Instructors also receive EUR 0.75 for each question they answer correctly, but may get an additional payoff depending on the treatment they are in. In the treatment *Fix*, instructors receive an additional payoff of EUR 4.50 no matter how good their pupil is. In the treatment *Linear*, instructors receive an

additional EUR 0.30 for each question their pupil answers correctly. In the treatment *Bonus*, they receive an additional EUR 9.00 if their pupil has at least 15 correct answers. For the treatment *Tournament*, the instructor-pupil pairs are each randomly assigned to groups of three pairs. The instructors then receive an additional EUR 13.50 if their pupil is the best one in their group. No show-up fee was paid in any of the treatments.

The amount of money given in the individual treatments was calibrated in such a way that the ex post payoff average for the instructors remains roughly constant in all treatments[6]. Otherwise, a difference in teaching between the treatments might be due to the higher amount of money earned and not the conditionality of the payoff.

In addition to the payoff scheme, a second treatment dimension is the subject pool used in the instructor role. To approximate actual teachers, students of educational science who want to become teachers are used[7]. In the following, the abbreviation "EDU" is used for those students, while "Non-EDU" designates other students.

Combining these two treatment dimensions results in 8 treatments. Table 1 shows the number of observations for each treatment.

TABLE 1: Number of observations by treatment

	Incentive Structure			
Subject Pool	Fix	Linear	Bonus	Tournament
EDU	24	23	22	22
Non-EDU	23	24	24	32

2.3 Procedures

The experiment was conducted computer-based and took place at the "Essen laboratory for experimental economics" (elfe) at the University of Duisburg-Essen in May and November 2011. Participants were recruited via the program ORSEE (Greiner, 2004) and the attached subject pool[8]. To program the experiment, the software z-Tree (Fischbacher, 2007) was used. For the video conference, a customized version of the software Vivicom was used. In total, 34 sessions with up to 12 subjects each were conducted, leading to a total of 392 participants[9]. The participants were all students from the University of Duisburg-Essen. To avoid a confounding of possible treatment effects with gender effects, all treatments included a balanced ratio of all possible gender pairings (male instructor/male pupil, female instructor/male pupil,

[6] This was done by first running the *Linear* treatment and then choosing the figures according to the results.

[7] The German system of teacher education is set up in a way that studying a subject to become a teacher is a distinct degree and university career from studying the subject per se (These students are called "Lehramtsstudenten"). For example becoming a chemistry teacher means studying "chemistry to become a teacher", and not the regular subject chemistry. This makes it possible to claim that a large share of the subjects used in the experiment will indeed become teachers.

[8] Table A in the appendix shows some descripitive statistics of the participants

[9] Of those, 4 subject pairs experienced technical problems and therefore are excluded from the analysis.

male instructor/female pupil, and female instructor/female pupil)[10]. The experiment lasted about 90 minutes for the instructors and 60 minutes for the pupils. Average payoff was EUR 20.23 (minimum EUR 9.75, maximum EUR 30.75) for the instructors and EUR 12.13 for the pupils (minimum EUR 4.50, maximum EUR 18.75). Subjects were paid out one after the other to preserve anonymity; the instructors were paid out before the pupils.

3 Hypotheses

Instructors in the present experiment can invest the restricted time during a session in learning and teaching. Even if both efforts are not completely separable it seems plausible to assume that instructors face a trade-off between teaching and learning due to both the restricted time available and the complexity of the content. From a theoretical point of view it is thus evident that homo oeconomicus like instructors do not exert teaching effort in the treatment *Fix*. In contrast, any of the implemented incentive schemes shifts priorities towards teaching under very weak assumptions[11]. If we define the share of knowledge an instructor is able to transfer to her pupil as "number of correct pupil answers" divided by "number of correct instructor answers" for each pupil-instructor pair, we can thus formulate hypothesis 1 as follows:

The lowest share of knowledge transmitted occurs in the treatment Fix.

This hypothesis is tentatively supported by previous research. Even though it is difficult to compare both the task as well as the incentive schemes in this paper to the previous literature, there is some evidence that monetary incentives can increase performance (e.g. Bull et al., 1987; van Dijk et al., 2001; Lazear, 2000; Shearer, 2004)[12]. These papers furthermore seem to point at better performance under tournament than under piece rate incentives. From a theoretical point of view, however, this result depends on preferences and expectations of agents and on the specific setup of the tournament[13].

As for the two different subject pools, a priori the rationale still holds: If teaching causes disutility, a monetary incentive for it will increase effort. While it can be argued that teachers are maybe intrinsically motivated and therefore should not react to monetary incentives, the before mentioned literature on teacher performance pay indicates that they do. This leads to hypothesis 2:

There is no difference in the incentive reaction between the subject pools EDU *and* Non-EDU.

[10] Due to the aforementioned exclusions, the balancing was not perfect.

[11] Of course, one can think of preferences or expectations that e.g. let instructors shy away from competition. These would need to be very strong, however, to prevent every kind of priority shifting through the incentive schemes.

[12] Note, that a lot of experimental studies discuss pay for performance that is added on top of a fixed wage. See Camerer and Hogarth for an overview. In contrast, participants in the treatment *Fix* of the current paper receive a higher fixed payoff than participants in other treatments. Against the background of the literature on efficiency wage theory and gift-exchange games (c.f. Akerlof, 1982; Akerlof and Yellen, 1990; Fehr et al., 1998) this aspect hinders a direct comparison.

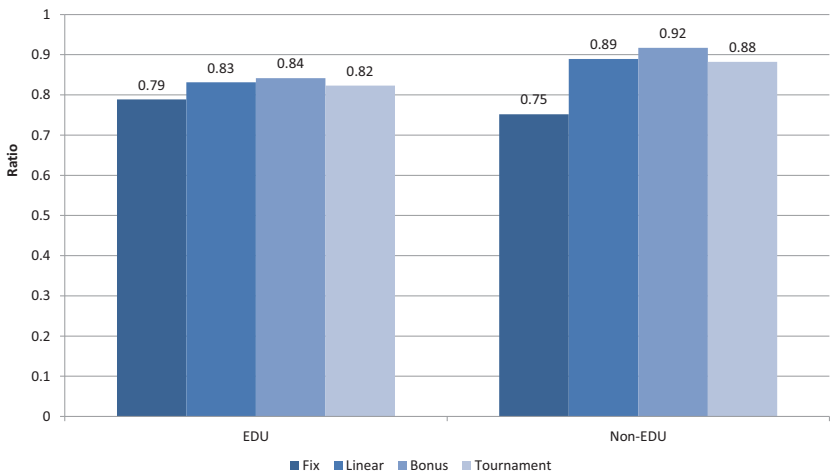
[13] See for example Harbring and Irlenbusch (2003) for an experimental investigation of the latter aspect.

4 Results

4.1 Treatment effects

Figure 1 shows the average share of knowledge transmission for all eight treatments. Looking first at the Non-EDU subjects, the highest average ratio is 0.92 in the *Bonus* treatment, followed by the *Linear* (0.89), the *Tournament* (0.88), and the *Fix* (0.75) treatment. Testing these differences pairwise reveals a significant difference between the *Fix* treatment and all other treatments ($p<0.05$ for all pairwise tests[14]). The treatments with a conditional monetary incentive (*Linear*, *Fix*, and *Tournament*) are not significantly different from each other in terms of the ratio (all tests yield $p>0.1$). Therefore, the formulated hypothesis cannot be rejected for the Non-EDU subjects.

FIGURE 1: Average ratio of correct answers pupil/teacher by treatment



Now looking at the EDU subjects, the ordering of treatments remains the same, but the difference between the *Fix* treatment and the other treatments is less pronounced: The highest ratio was achieved under the *Bonus* payment system (0.84), followed by the *Linear* (0.83), the *Tournament* (0.82), and the *Fix* (0.79) payment scheme. These differences are not significant (all tests yield $p>0.1$). Therefore, the hypothesis can be rejected for the subjects who study to become teachers: The payment scheme has no influence on their teaching performance[15].

Comparing the two different subject pools with each other for each treatment separately reveals that

[14] If not indicated differently, exact two-sided pairwise Wilcoxon rank-sum tests are used.

[15] Figures A and B in the appendix show the ratios of transferred knowledge for each of the two topics separately. Qualitatively, the results are the same as for both topics combined. The Non-EDU subjects have the lowest ratio of transferred knowledge in the *Fix* treatment. The difference between *Fix* and the other treatments is weakly significant ($p<0.1$) for the card game, while for the artificial language only the test of *Fix* vs. *Bonus* yields a weakly significant result. There are no systematic (or significant) differences for the EDU subjects.

there are no significant differences between prospective teachers (the EDU subjects) and students who do not want to become teachers (the Non-EDU subjects)[16]. Still, as shown previously, the reactions of the teachers on the different payment schemes are different for both groups.

FIGURE 2: Average number of correct teacher answers by treatment

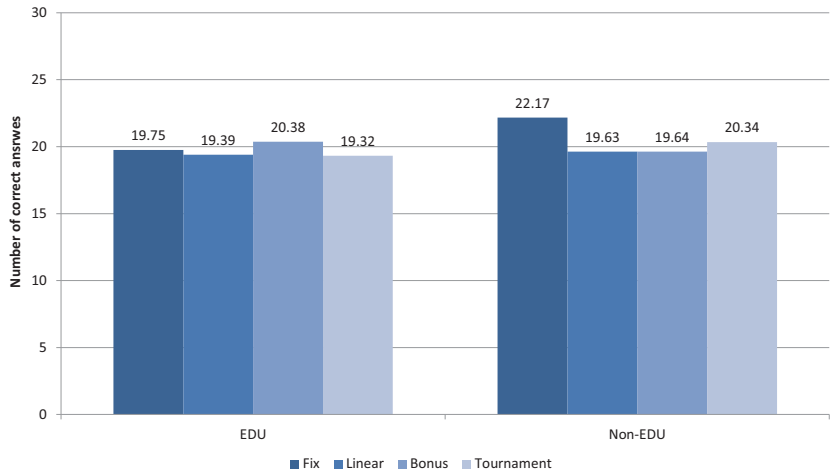
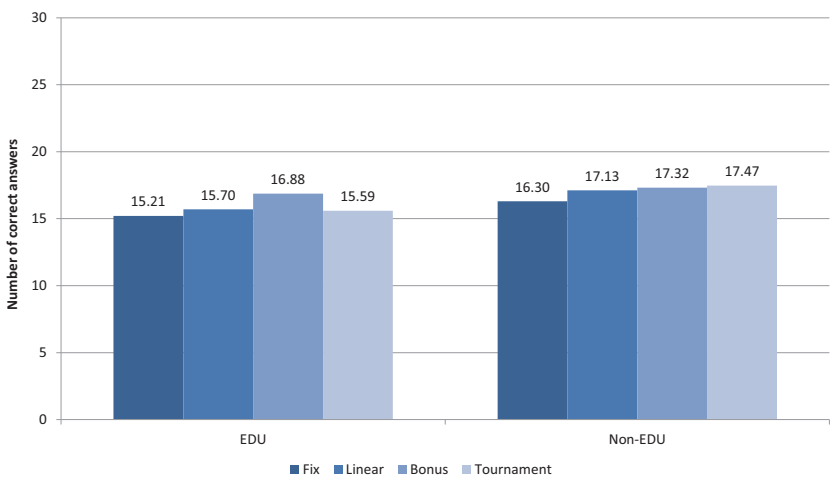


FIGURE 3: Average number of correct pupil answers by treatment



As the ratio of knowledge transfer is a combined measure of the correct answers by an instructor and her pupil, one can take a look at what drives the results in the ratio by analyzing these two underlying numbers shown in Figures 2 and 3. Again first looking at the Non-EDU subjects reveals that the difference

[16] All tests yield p-values > 0.1.

in ratio is driven mainly by the instructors' correct answers: The instructors have the highest number of correct answers in the *Fix* treatment[17]. The pupils have the lowest number of correct answers in the *Fix* treatment, this is however not significant[18]. One possible explanation for this result is that instructors have some limited capacity of effort which they can divide between two activities: Preparing to answer their own questions and preparing to teach their pupils. If the latter is incentivized, instructors increase their effort for this and therefore must decrease their effort for the preparation for their own questions, leading to the results presented here[19]. For the EDU-students, there are no significant differences for neither the instructors' nor the pupils' correct answers.

4.2 Video analysis

To open up the "black box" of the teaching process, the instructor videos are looked at directly and coded along several dimensions.[20] The analysis here is focused on the objectively codable variables: The time actually used to teach is recorded, starting with the first remark with content (thus excluding "Can you hear me?" or similar). Additionally, the time(s) when the topic is switched and the first topic are recorded, such that one can determine how much time was used for each topic. Furthermore, the type of pronoun used by the instructor (i.e. if the instructor refers to herself as "I" consistently or if she includes the pupil by using "we" at least once) and the type of address (i.e. if the instructor directly addresses the pupil at least once or if he uses an indirect address) are coded. Finally, it is recorded if the instructor points out to the pupil that she has the some material at her disposal.

Table 2 shows the instances of the variables "We" and "Direct Address" split along the two different subject pools. An exact χ^2-test shows that the difference between the instructors who study to become teachers and other instructors in their use of these two variables is statistically significant ($p<0.05$)[21]. A possible explanation for this behavior lies in the different perceptions prospective teachers might have concerning their role as instructors in the experiment. They see themselves as "proper" teachers and still try to keep their pupils engaged, leading to the use of a direct address. The other students however may not perceive themselves in a substantially different role than their pupils, thus using the pronoun "We" significantly more often[22].

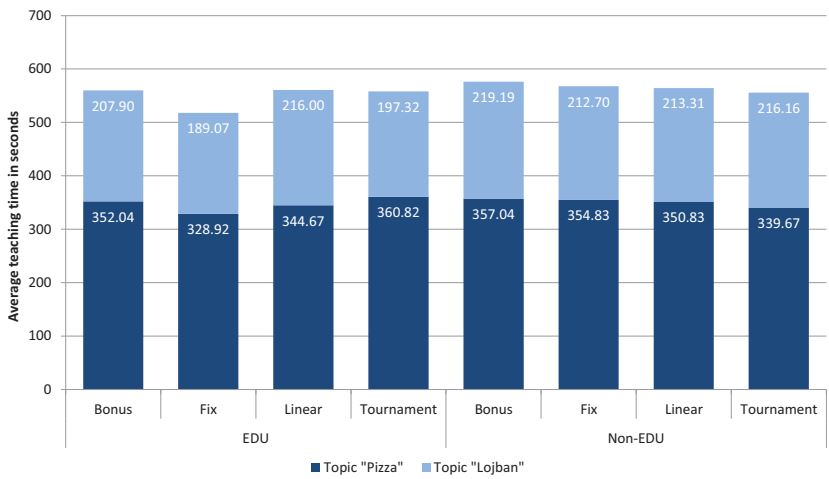
[17] $p<0.1$ for the comparison *Fix* vs. *Tournament*, $p<0.05$ for the other comparisons.
[18] All pairwise comparisons $p>0.1$.
[19] Note that the number of pupils' correct answers can be seen as a function of instructors' teaching effort and pupils' effort while the number of instructors' correct answers is only a function of their studying effort. Consequently, an effort shift because of changed incentives has stronger effects on the instructors' answers than on the pupils' answers if pupils' effort is assumed to remain constant.
[20] Table B in the appendix gives an overview of all coded variables. The more subjective variables are not used in the analysis. The coding of these variables proved to be too unreliable even though all coding was done by two different coders. This is especially true for the variables which measure how good each question is answerable with only the information given by the instructor. Of the 30 questions coded this way, in about one third the two coders had discrepancies more than 40% of the time.
[21] Comparing these variables along the different incentive schemes, however, does not reveal any systematic differences.
[22] Note that the type of address does not influence the results in terms of transferred knowledge.

TABLE 2: Use of direct address and we

	Direct address	We
EDU	68	60
Non-EDU	84	41
All subjects	152	101

Figure 4 shows the average time instructors use to teach the different topics. Instructors spent more time for the topic "Pizzabäcker" - which was also the one they saw first in the presentation - than for the topic "Lojban". However, there are no significant differences between treatments, regardless of the subject pool. The observed differences in outcome for the Non-EDU teachers are therefore not caused by simple quantity of teaching, but must have their causes in the manner of teaching.

FIGURE 4: Average time used to teach each topic by treatment



4.3 Regression

A different way to look at the data is by using a regression. Table 3 shows the results of a simple OLS regression with the share of transferred knowledge as the dependent variable. The regression is performed for the whole sample, just with the EDU students and just with the Non-EDU students.

The results partly confirm those from the non-parametric tests above: The treatment dummies for the *Linear* and *Bonus* treatment have a significantly positive effect (with the *Fix* treatment as a baseline) on the share of transferred knowledge only for the Non-EDU students. In terms of economic significance, the treatment effect for the Non-EDU students varies between 0.141 (*Bonus*) and 0.160 (*Linear*), so that the introduction of a monetary incentive increases the share of transferred knowledge by at least 14.1 percentage points. Comparing this to the average share of 75% in the *Fix* treatment for Non-EDU

TABLE 3: OLS estimation results for ratio of transferred knowledge

Dependent variable: Ratio of transferred knowledge	**All pairs**	**Only pairs with Non-EDU instructor**	**Only pairs with EDU instructor**
(Constant)	1.170***	1.562***	0.990**
	(0.000)	(0.002)	(0.034)
Linear treatment	0.164**	0.160**	0.039
	(0.019)	(0.033)	(0.575)
Bonus treatment	0.166**	0.141*	0.091
	(0.020)	(0.067)	(0.201)
Tournament treatment	0.100	0.095	0.025
	(0.128)	(0.168)	(0.724)
Linear x EDU	-0.132	–	–
	(0.168)		
Bonus x EDU	-0.090	–	–
	(0.355)		
Tournament x EDU	-0.072	–	–
	(0.457)		
Instructor EDU	0.016	–	–
	(0.823)		
Instructor female	0.011	0.008	0.019
	(0.752)	(0.866)	(0.716)
Pupil female	0.007	-0.002	0.057
	(0.837)	(0.971)	(0.315)
Teaching time	0.000	0.000	0.000
	(0.454)	(0.792)	(0.379)
Pupil material	-0.085**	-0.085	-0.071
	(0.022)	(0.130)	(0.188)
We or I	0.023	0.090	-0.057
	(0.519)	(0.103)	(0.290)
Direct Address	-0.055	-0.031	-0.065
	(0.205)	(0.674)	(0.244)
Card game experience instructor	-0.026	-0.047	-0.010
	(0.471)	(0.383)	(0.846)
Card game experience pupil	0.051	0.040	0.049
	(0.164)	(0.458)	(0.338)
Instructor's age	-0.006	-0.017**	0.006
	(0.284)	(0.040)	(0.431)
Pupil's age	-0.003	-0.009	-0.004
	(0.526)	(0.209)	(0.535)
Instructor's math grade	0.003	-0.015	0.023**
	(0.651)	(0.177)	(0.044)
Instructor's German grade	0.001	-0.004	0.010
	(0.888)	(0.794)	(0.438)
Pupil's math grade	0.000	0.000	-0.003
	(0.952)	(0.977)	(0.785)
Pupil's German grade	-0.007	-0.004	-0.031*
	(0.477)	(0.797)	(0.055)
Observations	180	92	88

Notes: Fixed treatment is the baseline for the treatment dummies. P-values in parentheses.
Significance levels: *10% **5% ***1%.
Source: Own calculations.

subjects shows that this is indeed a sizable increase.

In addition, other potentially confounding variables like age, sex, or school grades have no influence in the full sample. Finally, the only significant variable which deals with the way of teaching is the dummy if instructors have told their pupils that they have the same material at their place. This is somewhat surprising, as a priori the use of the material by the pupils should make it easier for them to understand the topics. However, a possible explanation for the decrease in transferred information might be that the instructor then omits some facts, thinking that the pupil can get them from the material.

5 Conclusion and discussion

Financial incentives conditional on the employee's performance are used in many, if not most, professional settings. In the realm of knowledge transfer, however, and especially in the education industry, they are not as widespread. This paper analyzes how different monetary incentives influence the effort and performance of people who are trying to transfer valuable knowledge: Students who want to become teachers are not influenced by different incentive schemes, while other students react in the expected way to monetary incentives: *Linear*, *Bonus*, and *Tournament* pay schemes all improve the ratio of transferred knowledge compared to a fixed wage. Between these three different schemes, there is however no difference. Looking at the time the subjects use to teach instead of the outcome, there is no difference between the different incentive schemes, regardless of the subject pool.

There are two possible explanations for the different results obtained with the two subject pools: People with certain preferences (such as high risk aversion or other-regarding preferences) might self-select into certain fields of study[23]. Dohmen and Falk (2010) show that this is the case for risk attitudes as more risk averse people select themselves into the teaching profession, thereby influencing the effect of certain incentive schemes. However, this can account for only a small part of the results in the paper at hand, as the *Linear* incentive scheme should be influenced by risk attitude less than the *Tournament* or the *Bonus* incentive schemes[24]. In addition, this paper's sample does not exhibit significant differences in risk or trust attitude as measured by a standard survey question in the questionnaire[25]. Furthermore, in most treatments neither trust nor risk attitude are significantly correlated with the ratio of transferred knowledge when tested for each treatment separately[26].

[23] Several studies find for example that students of economic subjects behave differently than students of other fields (Brosig-Koch et al. (2011); Ockenfels and Weimann (1999); Rubinstein (2006)). Brosig et al. (2010) find indication that these differences can be explained by selection rather than education effects

[24] In fact, the only influence of risk attitudes in the *Linear* scheme comes from the uncertainty how the student will perform given a certain teaching level, not from the incentive scheme itself.

[25] Dohmen et al. (2011) show that these survey questions are highly correlated with experimentally validated risk attitudes; Fehr et al. (2002) do the same for the trust question.

[26] This is done with Spearman correlation coefficients. The only significant correlations are between risk attitude and ratio of transferred knowledge in the *Fix* treatment with Non-EDU subjects ($p<0.05$) and in the *Linear* treatment with Non-EDU subjects ($p<0.1$).

A second and complementary explanation is that (prospective) teachers have a preference for the act of teaching itself. This could mean that such an intrinsic motivation would diminish the effect of an additional extrinsic source of motivation like money [27]. Such a crowding out effect would lead to no effect on teaching quality, matching the observed results. A more likely explanation could be that due to their intrinsic motivation, prospective teachers already give their maximum possible teaching effort, regardless of monetary incentives. This would mean that while introducing an additional monetary incentive might lead to the teachers wanting to increase their effort, this is just not possible. However, this would also mean that in the sample at hand prospective teachers have a lower ceiling in terms of teaching quality than the other students, as they are not able to transfer as much knowledge in the treatments with momentary incentives. An interesting follow-up study would be to look at the long term effect of monetary incentives on knowledge transfer, as the intrinsic motivation for teaching might change over time. This might help to entangle the two possible underlying causes described above.

[27] See for example Frey and Jegen (2001) or Frey (1997).

References

Adams, Scott J., John S. Heywood, and Richard Rothstein, *Teachers, Performance Pay, and Accountability: What Education Should Learn from Other Sectors*, Economic Policy Institute, May 2009.

Akerlof, George A., "Labor Contracts as Partial Gift Exchange," *The Quarterly Journal of Economics*, November 1982, *97* (4), 543–569.

_ **and Janet L. Yellen**, "The Fair Wage-Effort Hypothesis and Unemployment," *The Quarterly Journal of Economics*, May 1990, *105* (2), 255.

Brosig, Jeannette, Timo Heinrich, Thomas Riechmann, Ronnie Schoeb, and Joachim Weimann, "Laying off or Not? The Influence of Framing and Economics Education," *International Review of Economic Education*, 2010, *9* (1), 44–55.

Brosig-Koch, Jeannette, Christoph Helbach, Axel Ockenfels, and Joachim Weimann, "Still different after all these years: Solidarity behavior in East and West Germany," *Journal of Public Economics*, December 2011, *95* (11–12), 1373–1376.

Bull, Clive, Andrew Schotter, and Keith Weigelt, "Tournaments and Piece Rates: An Experimental Study," *Journal of Political Economy*, February 1987, *95* (1), 1–33.

Camerer, Colin F. and Robin M. Hogarth, "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," *Journal of Risk and Uncertainty*, 1999, *19* (1-3), 7–42.

Chaudhuri, Ananish, Sara Graziano, and Pushkar Maitra, "Social Learning and Norms in a Public Goods Experiment with Inter-Generational Advice," *Review of Economic Studies*, 2006, *73* (2), 357–380.

Dohmen, Thomas and Armin Falk, "You Get What You Pay For: Incentives and Selection in the Education System," *The Economic Journal*, August 2010, *120* (546), F256–F271.

_ , _ , **David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner**, "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," *Journal of the European Economic Association*, June 2011, *9* (3), 522–550.

Eberts, Randall, Kevin Hollenbeck, and Joe Stone, "Teacher Performance Incentives and Student Outcomes," *The Journal of Human Resources*, 2002, *37* (4), 913–927.

Fehr, Ernst, Erich Kirchler, Andreas Weichbold, and Simon Gächter, "When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets," *Journal of Labor Economics*, April 1998, *16* (2), 324–351.

_ , Urs Fischbacher, Bernhard Von Rosenbladt, Jürgen Schupp, and Gert G. Wagner, "A Nation-Wide Laboratory - Examining trust and trustworthiness by integrating behavioral experiments into representative surveys," *Schmollers Jahrbuch*, 2002, *122* (4), 519–542.

Figlio, David N. and Lawrence W. Kenny, "Individual teacher incentives and student performance," *Journal of Public Economics*, June 2007, *91* (5-6), 901–914.

Fischbacher, Urs, "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 2007, *10*, 171–178.

Frey, Bruno S., *Markt und Motivation. Wie ökonomische Anreize die (Arbeits-) Moral verdrängen*, München: Vahlen, 1997.

_ and Reto Jegen, "Motivation Crowding Theory," *Journal of Economic Surveys*, December 2001, *15* (5), 589–611.

Glewwe, Paul, Nauman Ilias, and Michael Kremer, "Teacher Incentives," *American Economic Journal: Applied Economics*, July 2010, *2* (3), 205–227.

Gneezy, Uri and Aldo Rustichini, "Pay Enough or Don't Pay at All," *Quarterly Journal of Economics*, 2000, *115* (3), 791–810.

Greiner, Ben, "An Online Recruitment System for Economic experiments," in Kurt Kremer and Volker Macho, eds., *Forschung und Wissenschaftliches Rechnen 2003. GWDG Bericht 63*, 2004, pp. 79–93.

Harbring, Christine and Bernd Irlenbusch, "An experimental study on tournament design," *Labour Economics*, 2003, *10* (4), 443–464.

Kessler, Ian and John Purcell, "Performance Related Pay: Objectives and Application," *Human Resource Management Journal*, March 1992, *2* (3), 16–33.

Lavy, Victor, "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement," *Journal of Political Economy*, December 2002, *110* (6), 1286–1317.

_ , "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics," *The American Economic Review*, December 2009, *99*, 1979–2021.

Lazear, Edward P., "Performance Pay and Productivity," *The American Economic Review*, December 2000, *90* (5), 1346–1361.

Martins, Pedro S., "Individual Teacher Incentives, Student Achievement and Grade Inflation," Technical Report, Queen Mary, University of London, School of Business and Management, Centre for Globalisation Research December 2009.

Muralidharan, Karthik and Venkatesh Sundararaman, "Teacher Performance Pay: Experimental Evidence from India," *The Journal of Political Economy*, February 2011, *119* (1), 39–77.

Ockenfels, Axel and Joachim Weimann, "Types and patterns: an experimental East-West-German comparison of cooperation and solidarity," *Journal of Public Economics*, February 1999, *71* (2), 275–287.

Pokorny, Kathrin, "Pay but do not pay too much: An experimental study on the impact of incentives," *Journal of Economic Behavior & Organization*, 2008, *66* (2), 251 – 264.

Prendergast, Canice, "The Provision of Incentives in Firms," *Journal of Economic Literature*, March 1999, *37*, 7–63.

Reback, Randall, "Teaching to the rating: School accountability and the distribution of student achievement," *Journal of Public Economics*, June 2008, *92* (5-6), 1394–1415.

Rubinstein, Ariel, "A Sceptic's Comment on the Study of Economics," *The Economic Journal*, March 2006, *116* (510), C1–C9.

Schotter, Andrew and Barry Sopher, "Advice and behavior in intergenerational ultimatum games: An experimental approach," *Games and Economic Behavior*, February 2007, *58* (2), 365–393.

Shearer, Bruce, "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment," *Review of Economic Studies*, April 2004, *71* (2), 513–534.

van Dijk, Frans, Joep Sonnemans, and Frans van Winden, "Incentive systems in a real effort experiment," *European Economic Review*, February 2001, *45* (2), 187–214.

Woessmann, Ludger, "Cross-country evidence on teacher performance pay," *Economics of Education Review*, June 2011, *30* (3), 404–418.

Appendix

Tables and figures

TABLE A: Descriptive statistics, individual level

Variable	Minimum	Maximum	Average	Standard deviation
Share of correct answers	0.27	1.43	0.818	0.209
Participants in the pupil role				
Age	19	42	23.920	3.953
Experience with card games (binary)	0	1	0.340	0.475
Experience with the game used	0	0	0	0
Experience with the language used	0	0	0	0
Teaching experience (binary)	0	1	0.510	0.501
Female (binary)	0	1	0.495	0.501
Last math grade	6	15	11.505	1.831
Last German grade	6	15	11.272	2.257
Risk attitude	0	10	5.540	2.233
Study time (in semesters)	1	23	4.980	3.779
Trust score	1	4	2.634	0.589
Number of correct answers	6	26	16.490	3.497
Total observations	194			
Participants in the instructor role				
Age	19	40	23.880	3.416
Experience with card games (binary)	0	1	0.371	0.484
Experience with the game used	0	1	0.010	0.101
Experience with the language used	0	1	0.010	0.101
Teaching experience (binary)	0	1	0.701	0.459
Female (binary)	0	1	0.526	0.501
Last math grade	6	15	11.622	1.924
Last German grade	6	15	10.778	2.236
Risk attitude	0	10	5.340	2.088
Study time (in semesters)	1	19	5.580	3.647
Last slide seen	21	43	37.890	4.784
Trust score	1	4	2.624	0.634
Number of correct answers	12	29	20.090	3.519
Total observations	194			

Source: Own calculations.

TABLE B: Video coding variables

Variable name	Description
Time total	Time (in seconds) the instructor uses to teach
Time Pizza	Time (in seconds) the instructor uses to teach the topic "Pizza Bäcker"
Time Lojban	Time (in seconds) the instructor uses to teach the topic "Lojban"
Number switches	Number of times the instructor switches the topic
First topic	Which topic does the instructor start with (0="Pizzabäcker", 1="Lojban")
Material pupil	Does the instructor indicate that the pupil also has the material? (0=no, 1=yes, 2=instructor is not sure)
Motivation	Does the instructor motivate the pupil? (0=no, 1=yes)
Payoff	Does the instructor talk about the experiment's payoff rule? (0=no, 1=yes)
Own effort	Does the instructor comment on her own teaching performance? (0=no, 1=yes)
Time teaching	Does the instructor mention the time constraint during teaching? (0=no, 1=yes)
Time presentation	Does the instructor mention the time constraint during the presentation? (0=no, 1=yes)
Summary	Does the instructor provide a summary for one or both of the topics? (0=no, 1=yes)
Both topics	Does the instructor mention both topics at the beginning of the teaching period? (0=no, 1=yes)
Cut off	Is the instructor cut off in the middle of teaching or does she finish by herself? (0=no, 1=yes)
We	Does the instructor use the pronoun "we" (or "us")? (0=no, 1=yes)
Direct address	Does the instructor adress the pupil directly? (0=no, 1=yes)
Misunderstood	Does the instructor misunderstand the situation (e.g. expecting the pupil to talk to him)? (0=no, 1=yes)
Comment	Does the instructor comment the situation (e.g. saying that it is strange not to hear the pupil)? (0=no, 1=yes)
Camera	Does the camera point at the instructor's face? (0=no, 1=yes)
Eye contact	Does the instructor establish eye contact with the pupil? (0=no, 1=yes)
Speed	How fast is the instructor's rate of speech? (0=slow, 1=average, 2=high)
Material "Pizzabäcker"	Does the instructor hold material (playing cards) in front of the camera? (0=no, 1=yes)
Examples "Pizzabäcker"	How many examples does the instructor use to explain the game?
Mistakes "Pizzabäcker"	How many mistakes does the instructor make when explaining the game?
Material "Lojban"	Does the instructor hold metarial (example sheet) in front of the camera? (0=no, 1=yes)
Examples "Lojban"	How many examples does the instructor use to explain the language?
Mistakes "Lojban"	How many mistakes does the instructor make when explaining the language?
Question n	How many answering possibilities from question n can the pupil exclude from the instructor's explanations? (if the pupil is led to believe the wrong answer is correct, this is coded as 4.) $n \in [1;30]$.

FIGURE A: Average ratio of transferred knowledge of the topic "card game" by treatment

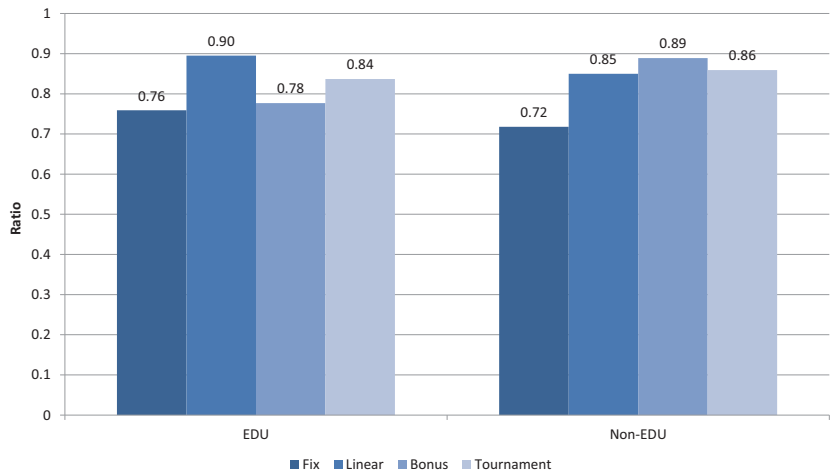
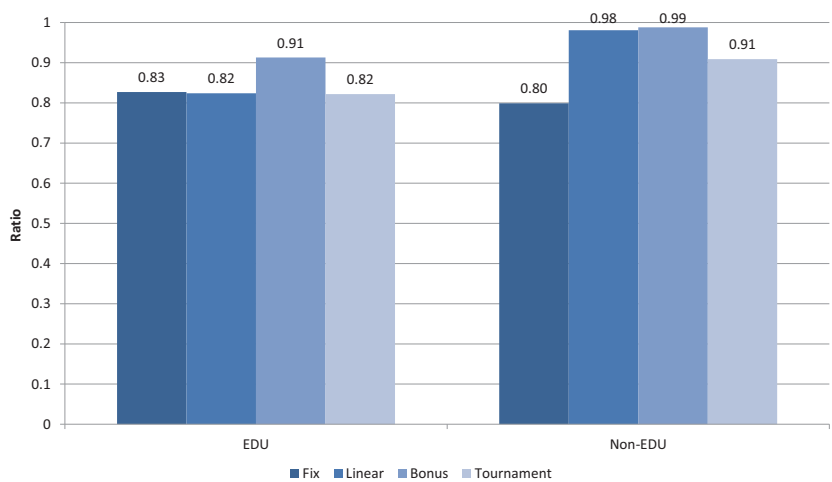


FIGURE B: Average ratio of transferred knowledge of the topic "artifical language" by treatment



Instructions[28]

Welcome to the experiment!

You are participating at a study of decision making behavior in the context of experimental economics. During the study you and the other participants will be asked to make decisions. You can earn money with this study. How much money you earn is dependent on the course of the experiment. You will receive detailed instructions about this in the following. All participants are paid in cash directly after the experiment one by one. To assure this, please remain seated after the experiment until your cabin number is called.

During the course of the experiment, no participant will receive information about the other participants' identity. All decisions are therefore made anonymously.

Should you have any questions before the start of the experiment, please ask an employee of the laboratory. He will come to your place and help you. **Any communication with the other participants during the experiment is only allowed when explicitly prompted; breaking this rule will lead to an immediate exclusion from the experiment.**

[28] These instructions are translated from the original, German instructions. Additional materials are available from the authors upon request.

Instructions

The experiment consists of three parts. In the following, you will receive detailed information about these. Please read the instructions carefully and thoroughly and click the start button on the screen **only after you have clarified all possible questions**. **After that, further questions cannot be answered any more.** In this experiment, you are either a **"teacher" or a "student". All participants receive the same instructions and materials, however.** At the beginning of the experiment, your role is displayed on the screen. You are allowed to make notes during the whole experiment. You are informed about the remaining time of the single parts on the computer screen.

Part 1

In the first part of the experiment, **teachers** see a **presentation** which contains **28 slides** for a total of **20 minutes**. In this presentation, two topics - on 14 slides each - are explained. Teachers can control the presentation with the buttons at the bottom of the screen. Some additional materials belonging to the two topics are included with the instructions. Students receive the same additional materials. After the 20 minutes are over, teachers have 10 minutes to prepare for part 2 of the experiment. In total, part 1 therefore lasts 30 minutes.

Students enter the laboratory during the course of the first part. They do **not** participate at part 1.

Part 2

Every teacher is randomly assigned one student. The computer automatically establishes a audio- and video-connection between a teacher and his student. Every teacher now has the chance to pass along the knowledge acquired in part 1 to his student. Picture and sound are transmitted only from the teacher. Each student can only confirm by clicking a button that he can see and hear the teacher at the beginning of the transmission. Before the transmission starts, teachers have one minute to adjust camera position and sound volume:

Camera position: Teachers can manually adjust the camera on the monitor in a way that they can see themselves in the middle of the right screen.

Sound volume: At the top of the video picture, teachers can as observe the volume level of their voice (dark blue area) when they speak into the microphone as seen on the following picture:

To adjust the volume level, a menu appears at the top of the screen when the mouse is moved there:

The indicator for the microphone (right) should be in a position that the volume level is clearly

reacting when speaking normally.

Additionally, students can adjust the volume directly at the headset's cord.

In general, the content of the communication can be chosen freely by the teachers. They are however not permitted to give personal information about themselves. This includes for example name, age, address, study subject, or similar. Breaking this rule will lead to exclusion from the experiment and therefore no payoff.

After 10 minutes, the transmission stops automatically and part 3 starts.

Part 3

All participants complete a multiple-choice test which contains questions related to both topics from the presentation in part 1. The test consists of 30 questions overall, 15 for each topic. There are four possible answers for each question, **exactly one** of those is correct. There are no deductions for wrong answers. You have 40 seconds for every question, so part 3 lasts 20 minutes in total.

Payoff

The **teachers'** payoff is the sum of two components:

1. Teachers receive **EUR 0.75** for each question they have answered correctly themselves.

2. Teachers receive a fixed amount of EUR 4.50. [Only in treatment *Fix*]

2. Teachers receive **EUR 0.30** for each question their student has answered correctly. [only in treatment *Linear*]

2. Teachers receive **EUR 9**, if their student has answered at least 15 questions correctly. [only in treatment *Bonus*]

2. Three randomly chosen teachers form a group. The teacher in this group whose student answered the most questions correctly receives EUR 13.5. In the case of a tie, the **EUR 13.50** are divided equally among the respective teachers. [only in treatment *Tournament*]

Students receive EUR 0.75 for every question they have answered correctly themselves.