WZB

Wissenschaftszentrum Berlin
für Sozialforschung

Armin Falk
Nora Szech

Pleasures of Skill and Moral Conduct

Discussion Paper

SP II 2016–301

January 2016

Research Area
Markets and Choice

Research Unit
Economics of Change

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Copyright remains with the authors.

Discussion papers of the WZB serve to disseminate the research results of work in progress prior to publication to encourage the exchange of ideas and academic debate. Inclusion of a paper in the discussion paper series does not constitute publication and should not limit publication in any other venue. The discussion papers published by the WZB represent the views of the respective author(s) and not of the institute as a whole.

Affiliation of the authors:

Armin Falk, Center for Economics and Neuroscience, University of Bonn
(armin.falk@uni-bonn.de)

Nora Szech, Department of Economics, Karlsruhe Institute of Technology,
Berlin Social Science Center (WZB), and CESifo (nora.szech@kit.edu)

Abstract

Pleasures of skill and moral conduct

As was recognized by Bentham[1], skillfulness is an important source of pleasure. Humans like achievement and to excel in tasks relevant to them [2, 3, 4, 5, 6, 7, 8, 9]. This paper provides controlled experimental evidence that striving for pleasures of skill can have negative moral consequences and causally reduce moral values. In the study, subjects perform an IQ-test. They know that each correctly solved question not only increases test performance but also the likelihood of moral transgression. In terms of self-image, this creates a trade-off between signaling excellence and immoral disposition. We contrast performance in the IQ-test to test scores in an otherwise identical test, which is, however, framed as a simple questionnaire with arguably lower self-relevance. We find that subjects perform significantly better in the IQ-test condition, and thus become more willing to support morally problematic consequences. Willingness to reduce test performance in order to behave more morally is significantly less pronounced in the IQ versus the more neutral context. The findings provide controlled and causal evidence that the desire to succeed in a challenging, self-relevant task has the potential to seduce subjects into immoral behaviors and to significantly decrease values attached to moral outcomes.

When you see something that is technically sweet, you go ahead and do it.
J. R. Oppenheimer, scientific leader of the Manhattan Project.

In science, perhaps the most striking example illustrating the tension between pleasures of skill and moral conflict considers the development of the atomic bomb. The Manhattan project was one of the most challenging engineering and scientific projects of the twentieth century, bringing together leading physicists of the time and providing scientists with a unique opportunity to excel[10]. Physicist Richard Feynman stated about the Manhattan Project: "… we *started* for a good reason, then you're working very hard to accomplish something and it's a pleasure, it's excitement. And you stop thinking, you know; you just *stop*"[11]. Even after knowing about the devastating power of the bomb, for "many of the physicists, the initial reaction was excitement over a spectacular demonstration of the successful 'technologically sweet' joint effort leading up to 'the gadget', as it was known" [12]. Another member of the Manhattan Project, Australian physicist Sir Mark Oliphant, commented that he "learned during the war that if … the work's exciting they'll work on anything", and continued that there is "no difficulty getting doctors to work on chemical warfare and physicists to work on nuclear warfare"[13]. However, after the dropping of the plutonium bomb on Nagasaki, which was deemed unjustified by many of the scientists, numerous members of the Manhattan Project started worrying about moral implications. Many suffered from feelings of depression, nausea, guilt and, in some cases, outright horror[13, 14, 15].

Conceptually, it has been argued that both the desire for signaling skillfulness and acting in accordance with moral values originate from the same source, a desire for positive self-image[16, 17, 18, 19]. This desire may stem from an uncertainty about who we are, how skillful, and how moral. If a task is morally difficult, but also considered as suitable for proving skillfulness, people may find it difficult to resist working on it. Our study investigates behavior in a decision context where the same activity simultaneously signals positive self-image with regard to one's skills, but negative self-image in the realm of morality, and reveals that this conflict bears the potential for moral transgression.

A paradigm for studying pleasures of skill and moral behavior. To identify potentially adverse effects of pleasures of skill on moral behavior, in our two main treatments we implemented a morally problematic task and varied whether or not subjects consider the task self-relevant in terms of skillfulness[20]. In order to confirm that the decision environment in our main treatments is of moral relevance to subjects, we ran two otherwise identical control conditions in a morally neutral environment.

In all four conditions subjects were administered the same task, a 52-item intelligence test. The test included items for fluid intelligence (Raven matrices) as well as crystallized intelligence (vocabulary test questions). Test scores were not incentivized. In all conditions subjects only received a show-up fee of 15 euros, i.e., payment was never performance dependent. At the end of the test, subjects were informed about their test score, together with the information whether they had performed above or below average, in comparison to other students who had taken part in the same test before. This average is based on about 1,700 observations from other students. Subjects knew that these students had participated in the test in previous and unrelated experiments.

To vary the self-relevance of the test we framed it either as an "IQ test" or simply as a "questionnaire". In the IQ-conditions (IQ) the instructions informed subjects that they are taking part in an IQ test, and that IQ is a crucial skill that is associated with important life outcomes such as wages or educational achievement. In contrast, in the No_IQ conditions subjects were told that they are taking part in a questionnaire study, without mentioning any connection to intelligence.

We chose an IQ test and varied the framing for two reasons. First, IQ is arguably one of the most important and universal personal characteristics in determining overall success in life[21, 22, 23]. For example, IQ is related to educational achievement, employment prospects, wealth acquisition and career outcomes. It is self-evident that people prefer a higher to a lower IQ and like to think of themselves (and have others think of them) as rather intelligent. The task therefore ensures high ecological validity in representing a measure of positive self-image with regard to skills. Second, using an IQ test allows us to easily vary the meaning and self-relevance of the task, by means of different framing, while sticking to the same task across all treatments. This is an essential feature of our experiment as it permits treatment comparisons relying on identical tasks[24].

Moral consequences were implemented using the Mouse Paradigm[25]. The paradigm involves the killing of mice, which is morally objectionable since killing an animal implies causing harm in an intentional and unjustified way. While there exists no universal consensus about how to define the content of morality, avoiding and preventing harm is a central element according to most notions of morality[26]. In contrast to choice paradigms that involve monetary payments, our choice context involves drastic and irreversible consequences[27]. In the Mouse conditions, each subject was "endowed" with a mouse. Subjects knew that their mouse is a young and healthy mouse, which will live in an appropriate, enriched environment, jointly with a few other mice, if it is not killed. To rule out any uncertainty about consequences, the instructions also informed subjects explicitly about the killing process: "The mouse is gassed. The gas flows slowly into the hermetically sealed cage. The gas leads to breathing arrest. At the point at which the mouse is not visibly breathing anymore, it remains in the cage for another 10 minutes. It will then be removed." As part of the instructions, subjects were also presented a picture of a mouse and a short demonstration video of the killing process[28]. In the video, four mice first move vividly in the cage, then they successively slow down. Eventually they die, with their hearts beating visibly heavy and slow.

It is important to stress that all mice used in the experiment are so-called "surplus" mice. They were bred for animal experiments, but turned out unsuited or unneeded for current animal studies. Such mice are perfectly healthy, but keeping them alive is costly. It is common practice in laboratories conducting animal experiments to gas surplus mice. Our study made it possible to rescue these animals. Surplus mice that were saved by subjects' behavior were purchased by the experimenters and are kept in enriched conditions, precisely as stated in the instructions. Thus, as a consequence of our experiment, many mice that would have otherwise died were saved.

Subjects in the Mouse conditions faced a morally problematic consequence of succeeding in the task: They were informed in the instructions that for each correctly answered question in the test, the likelihood that their mouse would be killed increased by 0.9 percentage points. This leads to a maximum killing probability of 46.8 percent, in case all questions are answered correctly. Subjects knew that the killing probability (between zero and 46.8 percent) resulting from their performance in the test was used to determine whether their mouse was killed or not. To allow

subjects to save their mouse with certainty, each of the 52 test items included the possible response option "don't know" (this holds for all four treatment conditions). This way subjects could answer each question incorrectly with certainty if they wanted to, independently of whether they knew the correct answers to the test questions.

In total we study four treatment conditions, see Table 1. In all four conditions the 52-item test was exactly the same. Possible experimenter demand effects, e.g., in the sense that subjects experienced pressure to complete the test, or had a desire to outperform former participants, are kept constant between conditions. The two main treatments IQ and NoIQ differ only in the framing of the test either as self-relevant in terms of skillfulness, or neutral. Our two control treatments vary the framing of the task in an identical way but implement a decision environment that is morally neutral. In both Control_IQ and Control_NoIQ, subjects just performed the task without any further consequences. A total of 301 undergraduate students were randomly assigned to one of the four treatment conditions (between subject design)[29].

Pleasures of Skill	Morality: Main treatments: Mouse	Morality: Control treatments: No Mouse
High	Treatment: IQ Framing: IQ test Morality: Killing of mice $N = 81$	Treatment: Control_IQ Framing: IQ test Morally neutral $N = 72$
Low	Treatment: NoIQ Framing: Questionnaire Morality: Killing of mice $N = 84$	Treatment: Control_NoIQ Framing: Questionnaire Morally neutral $N = 64$

Table 1. Design overview and treatment variations.

The main hypothesis to be tested was that pleasures of skill involve the potential to reduce moral behavior. In an environment in which pursuing pleasures of skill causes

negative moral consequences, decision makers face two competing sources of self-image[16, 18, 19]. Our treatments exogenously vary self-relevance of performance and therefore allow studying our question of interest in a tightly controlled way. Specifically, we hypothesized that subjects are more willing to sacrifice the life of a mouse in order to obtain a high and self-relevant "IQ" score rather than performing well in the more neutral "questionnaire" condition. Thus test scores should be higher in IQ than in NoIQ. To check whether subjects actually perceive task performance in our two main treatments as morally relevant we also compare test scores in IQ and NoIQ with those in Control_IQ and Control_NoIQ, respectively. Without morally problematic consequences, test scores should be higher in the latter two conditions, compared to the two main treatments. Furthmermore, the comparison of willingness to reduce test performance to save the mouse in IQ versus NoIQ allows to approximate the value of the mouse life per se. This moral value may be lower in IQ compared to NoIQ, if subjects causally lower their perceptions of what is a moral value in the IQ context.

Pleasures of skill undermine moral behavior. Figure 1 displays the cumulative distribution functions of total test scores for the two main treatment conditions. The figure shows a clear difference between test scores in the IQ condition (solid red line) and the NoIQ condition (solid black line). Mean test scores in IQ are 35.13 (std. dev. 12.09), compared to 28.78 (std. dev. 15.62) in NoIQ. Thus, relative to NoIQ, subjects reach about 22 percent higher test scores. This difference is statistically significant ($P<0.01$, $n=165$, two-sided rank sum test).

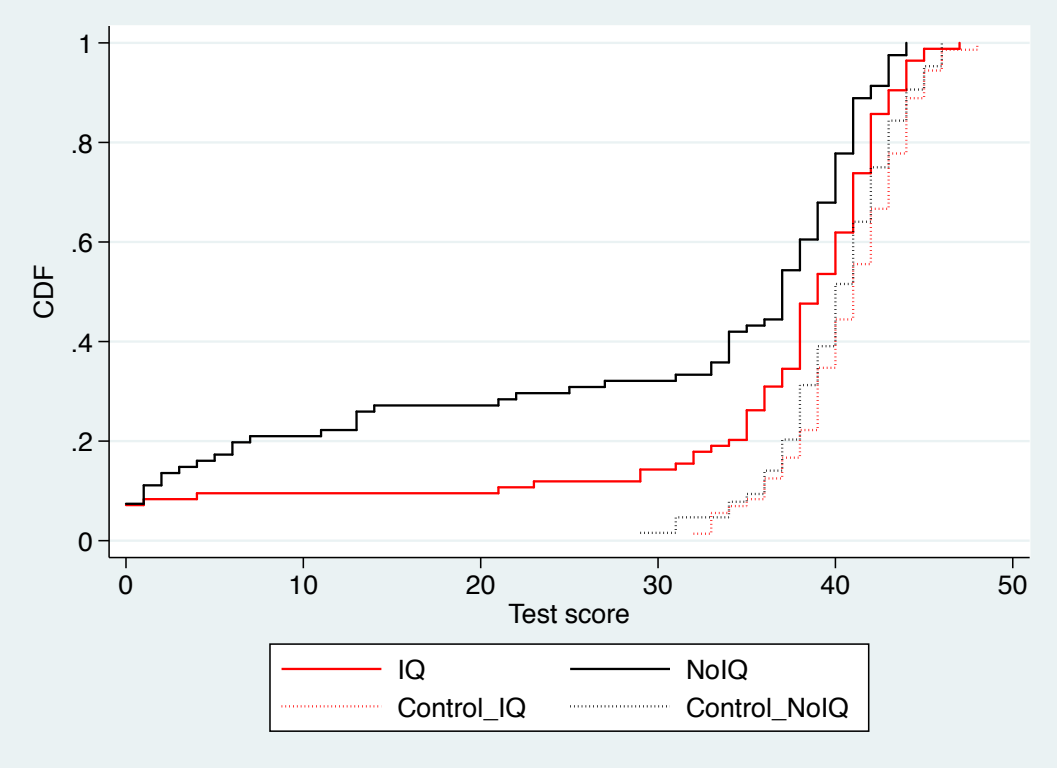

Fig. 1. Pleasures of skill undermine moral behavior. The figure displays cumulative distribution functions of total test scores for IQ (solid red line) and NoIQ (solid black line) as well as for the two control treatments (dotted lines). Test scores with negative moral consequences are significantly higher when framed as self-relevant in terms of skillfulness compared a neutral framing (P<0.01).

Comparing results from our control treatments with those of our main conditions further reveals that, as expected, adding a moral dimension reduces test scores. This holds for both the IQ conditions (P<0.01, n = 156, two-sided rank sum test) as well as the NoIQ conditions (P<0.01, n = 145, two-sided rank sum test). As can be inferred from Fig. 1, however, the decrease in performance is more pronounced in the NoIQ conditions than in the IQ conditions (P=0.013) *(30)*. This difference-in-differences result shows that when adding moral consequences, moral transgression is more likely in a self-relevant context, in comparison to a more neutral choice context. This suggests that subjects causally reduce their perceptions of what is a moral value "per se" in the IQ context. According to the smaller reduction in performance in order to save the mouse in the IQ context, the value of the mouse life is causally lower than in the NoIQ context.

A closer inspection of test score distributions across treatments shows that no subject scored below 29 in the morally neutral conditions. Thus 29 questions is a minimum bound of performance subjects are able to achieve in the test. In contrast, in

the two main conditions, a sizable fraction of subjects scored below this minimum level. In NoIQ 32.1 percent of subjects scored strictly below 29, significantly more than in Control_NoIQ (P<0.01, Two Sample Test of Proportion, two-sided, n=145). Thus while most of these subjects solved some test questions correctly, they apparently tried to ensure a relatively high likelihood of survival of their mouse and therefore reduced their test taking efforts. A possible interpretation is that subjects tried to find a compromise between revealing that they are capable of solving at least some of the questions and behaving morally. In the IQ condition such a compromise is rarely observed. Only 11.9 percent of subjects scored below the minimum level of 29, significantly less than in the No_IQ condition (P<0.01, Two Sample Test of Proportion, two-sided, n=165)[31]. This suggests that subjects in the IQ condition were keen to reach a rather high score placing relatively little weight on the life of their mouse. Among the few subjects who scored below 29, the majority solved exactly zero questions. Thus, these subjects ensured survival of their mouse with certainty, abstaining from pleasures of skill.

Discussion. We have shown that pursuing pleasures of skill in a self-relevant activity favors moral transgression in a causal and significant way. The temptation to excel provides a powerful motive to detach oneself from moral consequences. Compared to many real life situations, we believe that the decision context implemented in the experiment provides only a lower bound for the outcome relevance of this fundamental trade-off. First, in the Mouse conditions, moral consequences are transparent and unambiguous. Higher levels of performance immediately translate into higher killing probabilities. In addition, subjects were reminded of the consequences, saw the video about gassing mice and were told that mice that were saved would live in appropriate conditions. In contrast, in many real life scenarios actions and consequences are only indirectly linked, and consequences are often ambiguous. Second, task performance was not incentivized. Subjects received their show-up fee irrespective of performance and test score. They could easily save mice without any monetary loss. Third, the signal value of a high IQ score accrued only at the individual level; effects may be stronger if on top of pure self-signaling, monetary incentives, public recognition or reputation effects would become effective, in particular if moral consequences are ambiguous. In fact, in many real life applications material incentives, career or promotion concerns as well as public recognition are

characteristic consequences of high performance. Finally, the desire to signal skillfulness may be more pronounced in tasks, which are of higher self-relevance than succeeding in an IQ-test.

We believe that the trade-off between pleasures of skill and moral behavior, demonstrated in our study, is pervasive. People like to excel and enjoy pleasures of skill, regardless whether they are managers, bureaucrats, athletes or scientists. A problem arises if in doing a "good job", managers support firms producing morally questionable products, such as chemical weapons or landmines, bureaucrats serve terror regimes[32], athletes support cultures of doping[33], or scientists violate norms of good scientific practice[34], or develop research with potentially harmful implications, for example, because scientific inventions are misused. In these contexts, pursuing pleasures of skill exhibits the power to undermine morality.

Acknowledgements

We thank B. Enke, A. Greif-Winzrieth and J. Willrodt for excellent research assistance. We thank D. Bernheim, V. Crawford, T. Dohmen, L. Götte, J. Hermle, D. Huffman, G. Loewenstein, F. Rosar and N. Schweizer for helpful discussions. For technical, programming, or administrative support, we thank M. Antony, H. Gerhardt, and S. Hackel. A. Falk acknowledges financial support by the German Science Foundation (Deutsche Forschungsgemeinschaft) through the Leibniz Program. The mouse paradigm was approved by the Ethics Committee of the University of Bonn (reference no. 066/12).

References

1. Bentham, J. in *An introduction to the principles of morals and legislation.* (Oxford University Press, Oxford, 1907) chap. V.2.
2. The pleasure of mastery was also recognized by A. Marshall in his Principles of Economics[3]: "The importance of the mastery motive has been a large part of the demand for the highly skilled professional services and the best work of the mechanical artisan arises from the delight people have in the training of their faculties and in exercising them." See also the literature in psychology on the fundamental human need for "effectance", which describes the ability to successfully produce desired outcomes. For a recent account and references see [3, 4, 5, 6, 7, 8, 9].
3. Marshall, A. Principles of Economics, Digireads.com Publishing (2004).
4. Norton, N. et al., The IKEA Effect: When labor leads to love, Harvard Business School Marketing Unit Working Paper 11-091 (2011).
5. Ahuvia, A. Beyond the extended self: Loved objects and consumers' identity narratives, *Journal of consumer research 32(1): 171-184* (2005).
6. Belk, R. *Possessions and Self*, (John Wiley & Sons, New Jersey, 1988).
7. Dittmar, H. & Pepper, L. Materialistic values, relative wealth, and person perception: Social psychological belief systems of adolescents from different socio-economic backgrounds, Meaning, Measure and Morality of Materialism, Association for Consumer Research and School of Business, Provo, UT 40-45 (1992).
8. Furby, L. Understanding the psychology of possession and ownership: A personal memoir and an appraisal of our progress*, Journal of Social Behavior & Personality* (1991).
9. White, R. Motivation reconsidered: The concept of competence, Psychological review 66 (5), 297 (1959).
10. "The Los Alamos site was an exciting brewing pot of creative brains focusing on very difficult problems, involving both theoretical challenges, technological development, logistics and numerics." Ann-Marie Martensson-Pendrill. *Physics education* 41(6), p. 495 (2006).
11. Feynman, R. et al., *Surely You are Joking, Mr. Feynman! Adventures of a Curious Character*. (W. W. Norton, New York, 1985).
12. Martensson-Pendrill, A. The Manhattan project – a part of physics history. *Physics education* 41(6), p. 496 (2006).
13. Lifton, R. & Mitchell, G. *Hiroshima in America: Fifty Years of Denial.* (G. P. Putnam's Sons, New York, 1995) pp 31f. and 66.
14. For a general discussion see Russell, B. The social responsibilities of scientists, *Science* **131***: 391-392* (1960).
15. Weaver, W., Snow, C., Hesburg, T., Baker, W. The moral un-neutrality of science, *Science* **133***: 255-262* (1961).
16. Bénabou, R.& Tirole, J. Incentives and prosocial behavior, *American Economic Review,* 96(5), 1652-1678 (2006).

17. Bénabou, R. & Tirole, J. Self-confidence and personal motivation. *The Quarterly Journal of Economics* 117(3): 871-915 (2002).
18. Prelec, D. & Herrnstein, R. 13 Preferences or principles: Alternative guidelines for choice, *Strategy and choice* 319 (1991).
19. Loewenstein, G. Because it is there: The challenge of mountaineering - For utility theory, *Kyklos* (1999).
20. Details about the experiment and instructions are provided in SI.
21. Borghans, L. et al., The economics and psychology of personality traits, *Journal of Human Resources*, **43(4)**: 972-1059 (2008).
22. Sternberg, R. et al., The Predictive Value of IQ, *Merrill-Palmer Quarterly*, **47(1)**: 1-41 (2001).
23. The official taskforce of the American Psychological Association has defined intelligence as the "ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought" Neisser et al., Intelligence: Knowns and unknowns, *American psychologist* **51(2)**: 77 (1996).
24. Implementing physically different tasks, e.g., an IQ test versus a pointless transcription task, would pose the problem of dealing with different metrics in the dependent variable, and different production functions across tasks. This could bias findings in unknown ways, unrelated to our main hypothesis. Using the same task in all conditions, the dependent variable and effort cost functions for answering the test questions are always kept identical.
25. Falk, A. & Szech, N. Morals and Markets, *Science* **340**: 707-711 (2013).
26. Bernard, G. *The Definition of Morality, The Stanford Encyclopedia of Philosophy,* (ed. Zalta, E.) (2012) chap. V2.: "In this descriptive sense, although avoiding and preventing harm is common to all, "morality" can refer to codes of conduct of different societies with widely differing content, and still be used unambiguously."
27. In Deckers, T., Falk, A., Kosse, F. and Szech, N. Individual Characteristics as Source of Variation in Moral Outcomes, Working Paper, University of Bonn (2013) we provide convergent and discriminatory validity of the mouse paradigm as a measure for morality. Killing is negatively related to agreeableness, one of the Big-Five facets, which describes a tendency to be compassionate and cooperative rather than suspicious and antagonistic towards others; and positively related with Machiavellianism, measuring a person's tendency to be unemotional, and detached from conventional morality. Moreover, killing is not related at all to disposable income, whether students are professionally involved with animal research or animal experiments or a simple preference for animals, as expressed by having a pet at home.
28. See SI, Fig. S1.
29. See SI for details.
30. See Table S1 for details on the difference-in-differences estimation. Means and standard deviations of test scores in the control conditions are 40.7 (3.45) in Control_IQ, and 40.0 (3.64) in Control_NoIQ, respectively (see also Table S1).

31. This is also significantly different from the morally neutral Control_IQ condition ($P<0.01$, Two Sample Test of Proportion, two-sided, n=156).
32. In their documentary film "A specialist", Rony Brauman and Eyal Sivan portray Adolf Eichmann, a key figure in the organization of the Holocaust, as an expert who seeks to do a good job, taking pride in mastering a complex and technically challenging task. The task was to organize the deportation and ultimately the killing of millions of people. "A Specialist, portrait of a modern criminal"; a film by Rony Brauman and Eyal Sivan; Director: Leo T. Hurwitz (1999).
33. An example is Lance Armstrong who won the Tour de France seven times in a row. The titles were later withdrawn because of doping. In 2013, after years of denial, he admitted to having doped, though "not a lot" and with the intention to "level the playing field". (http://www.oprah.com/own/Why-Lance-Armstrong-Says-He-Had-to-Dope-to-Win-Video, 17 January 2013 and the videos therein.)
34. Two particularly striking examples of scientific misconduct are social psychologist Diederik Stapel who made up his data, as well as physicist Jan Hendrik Schön who duplicated data.

Supporting online Material

Pleasures of Skill and Moral Conduct

Armin Falk[1] and Nora Szech[2]

Overview Supporting Online Material

1. Additional table
2. Mouse figure and video
3. Procedural details
4. Instructions
5. References and notes

[1]Center for Economics and Neuroscience, University of Bonn, armin.falk@uni-bonn.de
[2]Chair of Political Economics, Institute of Economics, Karlsruhe Institute of Technology, Berlin Social Science Center, and CESifo Research Network Fellow, nora.szech@kit.edu

1. Additional table

Dependent variable: *Test score*	
Mouse dummy	-11.222***
	(1.795)
IQ dummy	.667
	(.611)
Mouse-IQ-Interaction	5.687**
	(2.266)
Constant	40.0***
	(.454)
Observations	301
R-squared	0.171

Table S1: OLS regression coefficient estimates, with test score as dependent variable and robust standard errors in parentheses. *Mouse dummy* is a dummy, which takes value 1 if observations come from mouse treatments. *IQ dummy* is a dummy, which takes value 1 if observations come from IQ treatments. *Mouse-IQ-Interaction* is an interaction of *Mouse dummy* and *IQ dummy*. *** indicate significance at the 1-percent-level, ** indicate significance at the 5-percent-level. The P-value of the Mouse-IQ-Interaction coefficient is 0.013. It may seem surprising that the difference between Control_IQ and Control_NoIQ is not more pronounced. Note, however, that if subjects in an experiment are asked to take part in a questionnaire, without an interesting alternative, there is no reason why they should not do so. In both control conditions test scores have no moral consequences, and subjects may have experienced an ambition to perform above average.

2. Mouse figure and video

Fig. S1. Picture of a mouse as presented in the instructions of the experiment[3]

Video: In the instructions subjects were informed about the process how mice are killed: "The mouse is gassed. The gas flows slowly into the hermetically sealed cage. The gas leads to breathing arrest. At the point at which the mouse is not visibly breathing anymore, it remains in the cage for another 10 minutes. It will then be removed." They were also shown a short video: The mice first move vividly in the cage, then they successively slow down. Eventually they die, with their hearts beating visibly heavy and slow. Comparable demonstration videos are publicly available at laboratories conducting animal experiments.

3. Procedural details. A total of 301 subjects (123 male, mean age = 20.19 (SD = 2.63) undergraduate university students from various majors, took part in the experiments. All treatments were computerized using z-Tree as experimental software *(1)*. Subjects were recruited using the software ORSEE *(2)*. The experiments were conducted in December 2013, at the Bonn Econ Lab. Subjects in all conditions only received a show-up fee of 15 euros.

To ensure credibility of our paradigm, we stressed right at the beginning that all statements made in the instructions are true, as is standard in economic

[3] This picture is licensed under the Creative Commons Attribution-Share Alike 1.0

experiments, and that all consequences of subjects' decisions are implemented exactly as stated.

Subjects were randomly assigned into treatment condition and every subject participated only in one condition. At the beginning of an experimental session, participants received detailed information about the rules and structure of the experiment. As part of the experiment, subjects answered a short questionnaire including personality (Big-5), socio-demographics and a few questions on general values (political and religious orientation and whether they are vegetarians) as well as on final high school grades.

4. Instructions

In the following we provide an English translation of the instructions for the main two conditions, IQ and NoIQ (see 4.1 and 4.2). The two control conditions were identical in wording, respectively, with the only difference that no mouse live was involved. Translations for the two other treatments are available upon request.

4.1 Instructions for the IQ treatment

Thank you very much for your participation! For your participation you will receive **15 euros**. At the end of the experiment you will receive your money in cash.

Please note: Throughout the whole experiment communication between the participants is not allowed. On the computer please only use the functions intended to be used. If you have questions please raise your hand. Your question will then be answered at your cubicle!

Please note: All statements made in these instructions are true. This holds for all experiments carried out by the Bonn Econ Lab, and also for this experiment. In particular, all actions to be taken will be implemented exactly in the way they are described.

15–item Big-5 survey

Test for the Measurement of Intelligence

In a moment you will be taking part in an exercise to measure your intelligence. The test that has been developed for this purpose consists of two parts that are both part of established techniques to measure the intelligence quotient (IQ). Both tests are standardized and frequently used in cognitive psychology.
Generally, intelligence is correlated with many factors of success in a person's life. These comprise, among other things, educational success and average life income.

The first part is a vocabulary test and measures the so-called crystalline intelligence. This test part has no time limit. The second part is a matrix test and measures the so-called fluid intelligence. This consists in turn of two phases, of which the second one has a time limit.

Both components are central devices to measure a person's overall intelligence.

In general: **The more questions you answer correctly, the higher is your measured intelligence quotient.**

The IQ-test you are going to work on in a moment has been conducted already with many students at the University of Bonn, in other experiments and as a part of a questionnaire without any further consequences. At the end of the IQ-test, we are going to tell you how well you performed in comparison to other students. To this end, you are going to learn how many questions on average students have previously answered correctly. You can then compare this value with your result. Your result in the questionnaire has another consequence for a mouse.

Details on the mouse

(Figure shown to subjects in the instructions.)

At issue is a healthy, young mouse. It lives with some other mice together in a small group. The expected lifetime of this mouse is approximately two years. The more questions you answer correctly, the higher is the probability that your mouse gets killed.

For each **correctly** answered question in the test, this probability increases by 0.9 percentage points. The maximal likelihood (if all questions are answered correctly) is 46.8 percent. Hence, if you, for example, answer 30 questions correctly, the probability is 27 percent; for 0 correctly answered questions, it is 0 percent; for 15 correctly answered questions, it is 13.5 percent; for 40 correctly answered questions, it is 36 percent etc.

If the probability is larger than 0 percent, a random generator determines according to this probability whether the mouse dies or lives.

Details on the killing process

If the random generator determines, according to this probability, that the mouse dies, the mouse is gassed. The gas flows slowly into the hermetically sealed cage. The gas leads to breathing arrest. As soon as the mouse is not visibly breathing anymore, it remains in the cage for another 10 minutes. It will then be removed.

Summary

In a moment you can start with the IQ-test. After finishing the tasks, you get to know whether your intelligence result is above or below the average of other students who participated in the test in other experiments. For this purpose, we will tell you your result and the average. The more questions you answer correctly, the higher is the probability that the mouse is killed.

Please press "Continue" when you have read the instructions. *(Button "Continue")*

Video

To visualize the killing of mice by gas, you will in the following see an excerpt of a documentation video (30 seconds). The mouse will be killed in an identical way.

Video

You can now start with the test. The more questions you answer correctly, the higher is the measured intelligence.

4.2 Instructions for the NoIQ treatment

Thank you very much for your participation! For your participation you will receive **15 euros**. At the end of the experiment you will receive your money in cash.

Please note: Throughout the whole experiment communication between the participants is not allowed. On the computer please only use the functions intended to be used. If you have questions please raise your hand. Your question will then be answered at your cubicle!

Please note: All statements made in these instructions are true. This holds for all experiments carried out by the Bonn Econ Lab, and also for this experiment. In particular, all actions to be taken will be implemented exactly in the way they are described.

15–item Big-5 survey

Test

In a moment you will be taking part in another questionnaire. The first part is about guessing words. The second part is about a hands-on assignment of symbols. The answering of the questions in the questionnaire has no further impact for you.

The questionnaire you are going to work on in a moment has been conducted already with many students at the University of Bonn, in other experiments and as a part of a questionnaire without any further consequences. At the end of the questionnaire, we are going to tell you how well you performed in comparison to other students. To this end, you are going to learn how many questions on average the students have previously answered correctly. You can then compare this value with your result. Your result in the questionnaire has further consequences for a mouse.

Details on the mouse

(See 4.1.)

Details on the killing process

(See 4.1.)

Summary

In a moment you can start with the questionnaire. After finishing the tasks, you get to know whether your result is above or below the average of other students who participated in the test in other experiments. To this end, we are going to tell you your result and the average. The more questions you answer correctly, the higher is the probability that the mouse is killed.

Please press “Continue” when you have read the instructions. *(Button “Continue”)*

Video

(See 4.1.)

Video

You can now start with the test.

5. References and notes

1. U. Fischbacher, z-Tree: Zurich Toolbox for ready-made economic experiments. *Experimental Economics* **10**, 171 (2007).

2. B. Greiner, An online recruitment system for economic experiments (in Forschung und wissenschaftliches Rechnen, GWDG Bericht 63, Ges. für Wiss. Datenverarbeitung, Göttingen, pp. 79–93 (2003).

Discussion Papers of the Research Area Markets and Choice 2016

Research Unit: **Economics of Change**

Armin Falk, Nora Szech

SP II 2016-301

Pleasures of skill and moral conduct

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice