



Entrepreneurial Discovery and the Demmert/Klein Experiment: Another Attempt at Creating the Proper Context

JANA KITZMANN
DIRK SCHIERECK

jana.kitzmann@ebs.de
dirk.schierECK@ebs.de

European Business School, International University, Schloss Reichartshausen, 65375 Oestrich-Winkel, Germany

Abstract. Friedrich Hayek conjectured that the free enterprise system is the most effective in making discoveries, and Israel Kirzner refines the conjecture by saying that profit opportunities evoke entrepreneurial discovery. Demmert and Klein (2003) present the first attempt to demonstrate the Hayek/Kirzner conjecture. On the whole, Demmert and Klein (2003) classify the results as disappointing but fruitful. In contrast we argue that additional experimental evidence might yield a demonstration of the conjecture. We continued the diligence and good-faith effort started by Demmert and Klein (2003) to devise an experimental setting that would create a genuine context for entrepreneurial discovery, yet the conjecture eludes our efforts at controlling the experiment. We duplicated the experiment at a Business School in Germany, with two simple variations. First, Demmert and Klein (2003) recruited only male students. We include male and female students. Second, Demmert and Klein (2003) used a payment schedule that includes a flat rate for participation and additional earnings depending on the presented performance. We drop the flat rate and slightly reduce the earnings per unit outperformance. Our results show that overall money matters. Our results are rather like those of Demmert and Klein (2003) and do not seem to be influenced by a baseline payment. Moreover, there are gender specific divergences showing male students earning significant higher additional earnings than their female fellow students.

Key Words: cross country divergences, entrepreneurial discovery, gender specific, monetary rewards

JEL classification: A1, D0

1. Introduction

Friedrich Hayek (1978) conjectured that the free enterprise system is the most effective in making discoveries. He emphasized the “role of knowledge and its enhancement through market interaction”,¹ hypothesizing that market participants, in order to approach an equilibrium state, would have to gain a better mutual understanding of their actions and plans. In other words, it must be explained by what *process* the market participants acquire the necessary knowledge.² This is predominantly the process of competition in a free enterprise system.³ Hayek regarded the latter as a discovery procedure, in which individuals would generate “such facts as, without resort to it, would not be known to anyone”.⁴ Kirzner (1985) refines Hayek’s conjecture by assuming that the availability of profit opportunities tends to evoke entrepreneurial discovery. He conjectured that the market participants’ entrepreneurial discovery would be evoked if they saw an opportunity for profit (“*alertness*”). The higher the degree of freedom in a society, the better it would be able to discover

social opportunities.⁵ Demmert and Klein (2003) present the first attempt to demonstrate the Hayek/Kirzner conjecture.

Previous experimental research on discovery processes, incentives and performance as a result of decisions has produced interesting results with regard to our study. Treating the aspect of insufficient information in economic decision situations, Pelham and Neter (1995) show that individuals often have no control over the context in which they take important judgments.⁶ In accordance with this finding, Sprinkle (2000) argues that incentive effects can not be found for single period decision making processes, like analysed by Pelham and Neter (1995), and rather suggests “multi period incentive contracts” leading to better decision results and thus better performance.⁷ Brown (1994) presents evidence that people learn from experience in economic settings and use reference points to take their decisions.⁸ Vandegrift and Brown (2003) analyse the fact that payoff gaps between winners and losers in a tournament situations do not necessarily increase performance. They come to the conclusion that for simple tasks a payment improvement does not lead to performance improvement, hence individuals with low capabilities will adopt strategies to enhance their performance in simple tasks.⁹

In 1999, Demmert and Klein performed an experiment with 135 male undergraduate students on their University campus in Santa Clara, California. They asked the students to transfer water from one bucket to another. The investigators’ aim was to demonstrate that entrepreneurial discovery was evoked by varying motivation and incentives. Four different and randomly assigned sub-samples of students were asked to perform an identical task, for which each received a different monetary reward. The variation in payoffs was supposed to test the hypothesis that stronger motivation would lead to a higher degree of entrepreneurial discovery.

We replicated the experiment at the European Business School in Oestrich-Winkel, Germany with slight variations. The most obvious difference in our experiment is that we tested not only male German students but also female students of business administration.

2. Experimental Design

2.1. *The Basic Experiment*

In the basic experiment of Demmert and Klein (2003), subjects were brought to the outdoor site where they were confronted with one full bucket of water and one empty bucket standing 15ft away. Next to the full bucket was a footstool with four plastic vessels of different sizes. The subjects could find printed instructions right next to the full bucket, where they were asked to transport as much water as possible to the empty bucket, without moving any bucket and only making one trip between the two buckets. The amount of payment according to the water transferred was indicated in the form of a payment schedule¹⁰ (see Table 2).

Our experiment followed the set-up of Demmert and Klein (2003). It took place on Tuesday, 5th May 2003. We were able to recruit 154 randomly assigned students for the experiment. The gender composition in comparison to Demmert and Klein (2003) was as follows (see Table 1).

Table 1. Gender composition of male/female undergraduate students in both experiments.

	Males	Females	Sum
European Business School (Kitzmann and Schiereck)	77%	23%	154
Santa Clara University (Demmert and Klein)	100%		135

The obvious method to transport the water is to use the four vessels and bring them to the empty bucket. The non-obvious and superior method is to fill with the vessels' water into the inverted footstool, fill the vessels and put them also into the inverted footstool, then to bring this to the empty bucket.

2.2. Preparation

Hence, we mainly pursued two goals when preparing the experiment:

- First, we wanted to avoid any kind of contamination of the subject population by prior knowledge about the experiment.
- Second, we tried to recreate exactly the characteristics of the Demmert and Klein (2003) experimental setup.

As for the first aspect, we had to run the four different sub-samples simultaneously. Otherwise, there would have been a high probability of an immediate spread of word about the experiment among the subjects. The best way to do so was during lecture time, where all subjects would be available (which gave us 90 minutes to perform the experiment). There were four parallel class sections in the 2nd semester (undergraduates), so that we decided to have four tester groups, each testing one of the four payment schedules.

Separate sites were selected on campus in order to avoid the subjects seeing one of the experimental sites when going out of their classrooms before actually being confronted with it themselves. We arrived at a total of eight sites, with each tester group consisting of three persons, i.e. one coordinator and two monitors.

Moreover, we repeatedly told the subjects not to speak about the experiment when returning to their fellow classmates (we briefed them via e-mail in advance, via in-class announcement before starting the experiment and via personal instruction after the experiment). The results indicate that contamination was avoided.

We tried to employ comparable equipment, most importantly the same kind of stool as being the crucial element of the experiment. In order to ensure comparability across the four sub-samples, all groups used ordinary coffee cups with jugs from the cafeteria, representing a slight difference to Demmert and Klein's (2003) design.

Furthermore, our group designed and transferred the following forms, questionnaires and instructions for uniform use across the four sub-sample groups:

The *On-Site Instructions* were analogous to those formulated by Demmert and Klein (2003). They contained information about the task of the experiment (transferring water),

the restrictions (only one trip, three minute time limit, no movement of the buckets, no conversation with the monitor) and the payment schedule.

We operationalized the measurement unit “cup” by “half a coffee cup”, so that the other groups (i.e. sub-samples 1A, 1B) could still use the seven unit threshold. For our payment schedule 2B, we paid 50 cents (1\$ = 1€) for each unit of water transferred, leading to following (expected) outcomes:

- Using the four coffee cups only = 8 units (4 €)
- Using the stool only = 24 units (12 €; exactly comparable to Demmert and Klein (2003))
- Using the stool and the cups = 32 units (16 €)

For last dry runs and coordination purposes, we met with all tester groups 2.5 hours prior to starting the experiment. We also used the time to prepare the sites and coordination points.

In the *Observation Form*, we collected general characteristics of the subjects (age, gender—as we used both male and female subjects—, respective semester group).

As in Demmert and Klein (2003), we had each subject complete an *Exit Questionnaire* after the experiment—one for those who did invert and another one for those who did not invert the stool. The first questionnaire asked whether the subject made a spontaneous or a deliberate discovery of the solution and whether he or she had any second thoughts on using the stool. The second asked whether the subject had thought about using the stool. If the answer was yes, the subject was asked why he or she did not use it. After a short in-class announcement, the procedure followed six steps:

1. The coordinator picks up four subjects in class;
2. The coordinator leaves two at designated waiting point (no view on experiment site);
3. The other two are guided to “coordination point”, fill out Observation Sheet and hand it over to monitor;
4. The subjects are individually picked up by monitors and taken to respective sites;
5. The subject is told to read Instruction Sheet carefully and perform experiment. During experiment monitor completes Observation Sheet;
6. The subjects return to coordination point, complete respective Exit Questionnaire, receive money and go back to class. Meanwhile, monitor cleans up the site and coordinator picks up the other two subjects from the waiting point/the next group of four from the class.

2.3. *Payment Schedule*

Like Demmert and Klein (2003), we settled on four different forms labelled 1A, 1B, 2A, and 2B, but varied the schedule in two ways (see Table 2). First, we dropped the baseline payment. Demmert and Klein (2003) offered \$12 in cash simply for participating. Because our experiments took place during the classroom lectures we did not need any monetary incentive to motivate the students for participation. Second, we slightly reduced the earnings per unit outperformance. Table 2 summarizes the payment schedule used and contrasts this to the original scheme.

Table 2. Comparison of payment schedules.

Form	Schedule for earning additional money	
	Demmert and Klein	Kitzmann and Schiereck
1A	If you do not transfer at least seven cups of water, you receive no additional money; if you do transfer at least seven cups of water, you receive an additional 10 cents for every cup transferred.	If you do not transfer at least seven cups of water, you receive no additional money; if you do transfer at least seven cups of water, you receive an additional 0.1€ for every cup transferred.
1B	If you do not transfer at least seven cups of water, you receive no additional money; if you do transfer at least seven cups of water, you receive an additional 1 dollar for every cup transferred.	If you do not transfer at least seven cups of water, you receive no additional money; if you do transfer at least seven cups of water, you receive an additional 0.2€ for every cup transferred.
2A	Transfer as much water as you can; [i.e. this form provided a zero marginal payoff for additional water transferred].	Transfer as much water as you can; [I.E. this form provided a zero marginal payoff for additional water transferred].
2B	You will receive an additional 50 cents for every cup of water transferred.	You will receive an additional 0.5€ for every cup of water transferred.

3. Results

3.1. Analysis

At the European Business School, during the 90 minutes time frame, a total of 154 undergraduate students (77% male and 23% female) were put through the experiment on 8 parallel test sites (see Table 1). The subjects were allocated to the sub-samples as follows: 42 subjects had to perform experiment subtype 1A, 40 subjects went through sub-type 1B, 37 subjects type 2A and 35 subjects sub-type 2B.

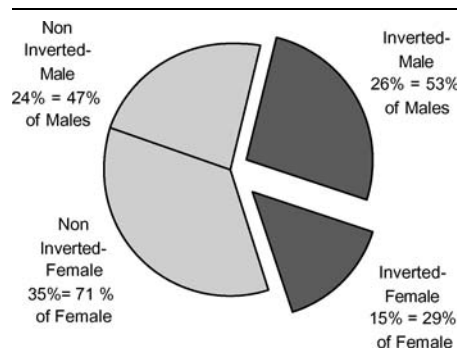
The results are rather similar across the two experiments (see Table 3). The baseline payment obviously does not improve results. However there are gender specific divergences that are remarkable (see Table 4). 53% of the male participants inverted the stool as opposed to only 29% of the females.¹¹ Using the four-part breakdown, the gender difference is not statistically significant difference (Chi-square = 1.94; p -value = 0.16). However, by testing for the dependency of gender and inverting the stool for the whole sample (i.e. the 154 observations), the difference becomes statistically significant (Chi-Square = 4.33; p -value = 0.04).

At the European Business School the portion which inverted the stool accounted for 40.9% (= 63 subjects) of the whole sample in contrast to only 31.1% in the U.S. sample. To perform

Table 3. Santa Clara University students vs. European Business School students.

Reward US\$/€	Santa Clara University					European Business School				
	Inverted	Not inverted	Total	Inverted	Not inverted	Total	Inverted	Not inverted	Total	Total
1A 0.1\$/€	11	26.19%	24	25.81%	35	12	19.05%	30	32.97%	42
1B 1\$/0.2€	13	30.95%	24	25.81%	37	22	34.92%	18	19.78%	40
2A 0	8	19.05%	24	25.81%	32	14	22.22%	23	25.27%	37
2B 0.5\$/€	10	23.81%	21	22.58%	31	15	23.81%	20	21.98%	35
	42	100.00%	93	100.00%	135	63	100.00%	91	100.00%	154

Table 4. Inversion of stool and gender dependence.



a Chi-Test, we classified the whole sample along the four sub-sample types and “Inverted Stool vs. Not-Inverted Stool”, resulting in 8 possible classifications. The Chi-Test leads to the conclusion that the differences are not statistically significant (Chi-Square-Test = 5.46; p -value: 0.14). Hence, we cannot reject the hypothesis that “sub-sample type” and “inverting the stool” are independent. After merging the categories 2A with 1A (relatively low incentive) and 2B with 1B (relatively high incentive), we repeated the Chi-Square-Test for independence with the now clustered categories: In the “low incentive” group (2A + 1A), only 33.8% inverted the stool, whereas 48.6% used this superior transport method in the “high incentive” group (1B+2B). This difference is significant on the $\alpha = 0.1$ level (Chi-Square = 3.53; p -value = 0.06), indicating that being in the “high incentive” group had a positive influence on the probability of inverting the stool. This supports the original test hypothesis that a higher monetary incentive would lead to a higher degree of entrepreneurial discovery.

Although our analyses show some differences between the European Business School students and the Santa Clara University students, we would not interpret this as cross-country divergences. The samples are too small and not finally random.¹² Similar to our sub-sample 2B at the European Business School, the portion of subjects who did invert the stool due to a spontaneous entrepreneurial idea was relatively small (22.2%), leaving

the rest to “problem solving” behavior. This means that the experiment might not be able to test for entrepreneurial discovery at all, hence problem solving in the context of this paper differs from entrepreneurial discovery/entrepreneurship. According to Kirzner (1979), entrepreneurship is “what sets [. . .] apart from knowledge as a resource that is, from problem solving, is reflected in one’s lack of self-consciousness concerning it.” Consistent with this, entrepreneurial discovery is undeliberate but motivated leading to a fundamental interpretive shift, whereas problem solving behavior is deliberate and leading to a partial interpretive shift.

The Exit Questionnaires show that not one of the 154 subjects knew the experiment beforehand, confirming that the sample population was not “contaminated”.

3.2. *Discussion of Results*

There are indications for differences in gender as for divergences relating to the incentive scheme. The fact that the high incentive group inverted the stool more often than the low incentive group is in the direction of the conjecture, but should still be interpreted with caution.¹³ Several factors might distort this already statistically weak result.

Demmert and Klein (2003) doubt that the experiment is really able to operationalize large enough variations in motivation. This is due to the potential presence of non-monetary rewards leading to even lower perceived variation of rewards. In our experiment, this flaw might even be stronger because the marginal reward for sub-sample 1B was changed from 1\$ to 0.2€.

The Exit Questionnaires show that among the European Business School subjects who did not invert the stool, 38.5% thought about doing so.¹⁴ Of the subjects that indicated that they thought about inverting the stool, 51.4% where in the high-incentive sub-samples (i.e. sub-sample Type 1B and 2B).¹⁵ It seems, consequently, that differences in monetary rewards do not explain why the subjects did not realize their idea. This could be explained by the fact that they did not perceive the situation as a worthwhile profit-opportunity (Compare Demmert and Klein 2003:308). Hence, either problems in estimating the advantage of using the stool or being unsure about the legality of using it have most probably led to the high portion of subjects rejecting their initial idea. As a result, it is doubtful that the experiment succeeded in creating a context for entrepreneurial discovery.

Summarizing the discussion, we made a diligent, good faith effort to use experimental methods to demonstrate the Hayek/Kirzner conjecture. The exercise did not fully succeed. Now, there are two ways to think about the experience. First, that the conjecture failed. Second, that the experimental methods failed. There were at least three major reasons for saying that the experimental effort—not the conjecture—failed:

1. There can be strong motivational factors aside from the monetary payoffs. These are also explained in the Demmert and Klein (2003) paper on page 307.
2. 38.5% of those who did not invert the stool said in the exit questionnaire that they had thought of doing so. So, obviously failing to perceive the idea was not what was standing in the way. Apparently, it was not an opportunity, since they perceived the idea and did not find it worth doing.

3. 78.8% of those who did invert the stool said they deliberately searched for the best way to transfer water (as opposed to the idea of inverting just coming to them). It would seem that discovery based on such deliberate search is not entrepreneurial, but rather problem solving, not unlike deliberately solving a puzzle. We do not call people entrepreneurial for just being bright.

So, although the experimental methods were diligently pursued, they were not able to test the conjecture properly.

4. Conclusion

We gave the experiment another try, hoping that perhaps the confounds (as listed in 1, 2, 3 above) would be less present and hence that we could have more faith in the exercise as a bona fide test of the conjecture. Unfortunately, we have little faith that our exercise was any better than the original attempt, as in our case 38.5% of non-inverters said they thought of doing so, and 78.8% of the inverters said they deliberately searched for the idea.

We extend the experimental design developed in Demmert and Klein (2003) to integrate female students and to vary the payment schedule. This allowed us to for gender specifica in entrepreneurial discovery, as far as it can be tested with this experiment. Our results provide evidence that males are more able in using opportunities. Male subjects tend to seize the opportunity of using the stool more often than female subjects. In comparison to the Santa Clara University sample of Demmert and Klein (2003), we can argue that the subjects at European Business School differ with regards to their ability to perceive undeliberate and nonobvious opportunities.

As we encountered problems similar to those of Demmert and Klein (2003), one could make an argument for possible improvements of the experimental design: we would suggest testing for further “experience” variables (such as age or prior job experience) and their influence on entrepreneurial discovery rather than slightly varying monetary rewards, whose impact on evoking entrepreneurial discovery could not be proven. Toward this end, including variations of non-monetary rewards, e.g. a publicly announced reward for the “winning subject” (i.e. the one who would transport most water), can enhance the experiment. Further research can elaborate on the observed gender differences or influences of other variables as planned majors or jobs, i.e. future interests of the subjects.

However, the more fundamental problem is to effectively test for entrepreneurial discovery at all! As a result of the post-experiment questionnaires, we find that 78.8% of the subjects who inverted the stool acted from a “problem solving” perspective. This highlights the aspect that the subjects “who inverted the stool were ‘thinking outside the box’, but in as much as they set out to solve a conceptual problem, they worked within a larger, more fundamental box, within which the subject’s deliberate thoughts remained” (Demmert and Klein 2003:308). Such deliberate problem solving does not qualify as Kirznerian discovery.

Consequently, future research has to look for other, potentially less formal ways to examine whether entrepreneurial, i.e. undeliberate and nonobvious discovery, can be evoked by setting higher monetary incentives in order to highlight its role in economic systems.

Overall, the results support the notion that money matters. A higher degree of monetary incentives is associated with a higher probability of discovery.

Acknowledgment

We are grateful for comments received from participants at the 2004 Annual Conference of the International Academy of Business & Public Administration Disciplines. We also thank Dan Klein and an anonymous referee for their specific comments and recommendations. Finally, we have to thank the participants of the 2003 “Testing for Rationality” seminar at the European Business School. Of course, all remaining errors are our own.

Notes

1. Kirzner (1997:67).
2. Compare Hayek (1948:46).
3. Compare Kirzner (1997:69).
4. Hayek (1978:179).
5. Compare Demmert and Klein (2003:297).
6. Compare Pelham and Neter (1995:592).
7. Compare Sprinkle (2000:303).
8. Compare Brown (1994:396).
9. Compare Vandegrift and Brown (2003:494).
10. The full set of instructions is available upon request from the authors. We would like to thank Patrick Heine-
mann, Jan Lennertz and Benjamin Moldenhauer for their support.
11. Gneezy, Niederle and Rustichini (2003) report significant gender gaps in mixed tournaments and better
performance of women in single sex tournaments (compare p. 1070). They analyze gender differences in
reaction to competitive incentive schemes, which is comparable to this study, and came to the conclusion that
in mixed groups the male participants’ performance rose significantly in reaction to performance variations
whereas the females’ performance did not. In an only female competitive environment, the mean performance
of the women rose within their group as well as in relation to the mixed group.
12. The results ($\text{Chi-Square} = 2.99$; $p\text{-value} = 0.08$) have a considerably lower $p\text{-value}$ than the sub-sample
comparison and the difference is significant at $\alpha = 0.1$. At European Business School, those class groups are
selected on the basis of entry examinations.
13. The sorting of the student body into the four groups was assignment randomly by the university.
14. Compare Nisbett and Wilson (1977) for a discussion of an individuals’ introspective access to cognitive
processes. It could be argued in this context that the subjects did not consciously think about doing so.
15. One could conclude that the effect of monetary incentives on whether subjects realize their idea will depend
on the comparison of “did it” vs. “thought about it and did not do it.”

References

- Brown, P. (1995) “Learning From Experience, Reference Points and Decision Costs.” *Journal of Economic Behavior and Organization*, 27: 381–399.
- Demmert, H. and Klein, D. B. (2003) “Experiment on Entrepreneurial Discovery: An Attempt to Demonstrate the Conjecture of Hayek and Kirzner.” *Journal of Economic Behaviour & Organization*, 50: 295–310.
- Gneezy, U., Niederle, M., and Rustichini, A. (2003) “Performance in Competitive Environments: Gender Differences.” *Quarterly Journal of Economics*, 118: 1049–1074.
- Hayek, F. A. (1978) *Competition as a Discovery Procedure*, In *his New Studies in Philosophy, Politics, Economics and the History of Ideas*. Chicago: University of Chicago Press, pp. 179–190.

- Kirzner, I. M. (1979) *Perception, Opportunity, and Profit: Studies in the Theory of Entrepreneurship* Chicago: University of Chicago Press.
- Kirzner, I. M. (1985) *Discovery and the Capitalist Process*. Chicago: University of Chicago Press.
- Nisbett, R. and Wilson, T. (1977) "Telling More than We Can Know." *Psychological Review*, 84: 231–259.
- Pelham, B. W. and Neter, E. (1995) "The Effect of Motivation on Judgment Depends on the Difficulty of the Judgment." *Journal of Personality and Social Psychology*, 68: 581–594.
- Sprinkle, G. B. (2000) "The Effect of Incentive Contracts on Learning and Performance." *The Accounting Review*, 75: 299–326.
- Vandegrift, D. and Brown, P. (2003) "Task Difficulty, Incentive Effects, and the Selection of High-Variance Strategies: An Experimental Examination of Tournament Behavior." *Labour Economics*, 10: 481–497.