

**ON COOPERATION AND TRUST
IN STRATEGIC GAMES**

Behavioral Evidence from the Middle East and Europe

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften

durch die

Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität zu Bonn

vorgelegt von:

GARI WALKOWITZ

aus Berlin

Bonn, 29. März 2010

Referent: Prof. Dr. Dr. h.c. mult. REINHARD SELTEN
Laboratory for Experimental Economics, University of Bonn

Ko-referent: Prof. Dr. ARMIN FALK
Institute for Empirical Research in Economics, University of Bonn

Dekan: Prof. Dr. CHRISTIAN HILLGRUBER

Tag der mündlichen Prüfung: 29. März 2010

Contents

Introductory comments	1
Chapter I: On the Prevalence of Framing Effects Across Subject Pools	5
I.1 Introduction	5
I.2 Experimental framework: Two cooperation games	8
I.2.1 Continuous prisoner’s dilemma with positive externality (PDP)	8
I.2.2 Continuous prisoner’s dilemma with negative externality (PDN)	9
I.2.3 Equivalence of the two games	9
I.3 Experimental procedures	11
I.4 Results	13
I.4.1 General cooperation in different subject pools	13
I.4.2 The impact of framing in different subject pools	13
I.4.3 Treatment-dependent cooperation across subject pools	15
I.5 Summary and Discussion	16
I. References	19
I. Appendices	23
Chapter II: The Moderating Effect of Conformism Values on the Relations between Other Personal Values, Social Norms, Moral Obligation, and In- dividual Cooperative Behavior	28
II.1 Introduction	28
II.2 Study 1: Conformism Values and Social Norms	33
II.2.1 Method	33
II.2.2 Results and Discussion	35
II.3 Study 2: Conformism Values and Cooperative Behavior	36
II.3.1 Method	37
II.3.2 Results and Discussion	38
II.4 Study 3: Mediation of Moderating Effects	40
II.4.1 Method	42
II.4.2 Results and Discussion	43
II.5 General Discussion	46
II. References	50
II. Appendix	55

Chapter III: Actions and Beliefs in a Trilateral Trust Game Involving Germans, Israelis, and Palestinians **56**

III.1 Introduction 56

III.2 Research questions 60

III.3 Experimental Methods and Procedure 61

 III.3.1 Experimental methods 61

 III.3.2 Experimental procedure 63

III.4 Results 64

 III.4.1 The impact of culture on senders’ trust and responders’ trust beliefs . 65

 III.4.2 The impact of culture on reciprocity and senders’ reciprocity beliefs . 70

 III.4.3 Return on investment 77

III.5 Discussion 78

III.6 Conclusion 81

III. References 82

III. Appendices 88

Chapter IV: Trust and Discrimination in the Labor Market - An Experimental Study with Criminal Offenders Appendix **103**

IV.1 Introduction 103

IV.2 Ex-offenders and the labor market - related literature 105

IV.3 An experimental approach 107

 IV.3.1 Why a laboratory experiment? 108

 IV.3.2 Methodological challenges and causal inference 109

 IV.3.3 Experimental design 110

 IV.3.4 Treatment variables 111

 IV.3.5 Experimental procedure 112

 IV.3.6 Behavioral prediction 115

IV.4 Results 116

 IV.4.1 Hiring (rank order) 117

 IV.4.2 Wage payment (transfer behavior) 120

 IV.4.3 Exerted efforts (reciprocity) 123

 IV.4.4 Do specific groups of employees face discrimination? 124

IV.5 Conclusion 132

IV. References 135

IV. Appendices 139

Introductory comments

This work is about selected issues of cooperation and trust in strategic games. Both somewhat abstract terms represent core mechanisms for decision making processes in uncertain or risky environments (e.g., political negotiations between conflicting parties, foreign direct investments, employment considerations). Cooperation literally means “working together” which can be seen as the root of all modern uses of the term - it denotes activity conducted between two or more people, in order to reach a common goal. Trust, on the other hand, is a motivation conditioned on cognitive and contingently emotional processes to take the risk of vulnerability via the achievement of an investment without explicit means of control within exchange relationships. Trust can be considered as a specific attitude to cooperate. Combining the intention to cooperate and the ability to trust and to positively reciprocate creates cooperation gains for all members of a social system (e.g., Israelis and Palestinians working together and sharing the same neighborhood) and generates valuable social capital. As FUKUYAMA (1995)¹ and KNACK and KEEFER (1997)² point out, higher trust is associated with stronger economic performance as well as with higher and more equally distributed incomes.

The present thesis is a result of my research during the last five years at the Laboratory for Experimental Economics (BonnEconLab) at the University of Bonn and other institutions worldwide. It is particular because it bunches subjects and methodologies across several borders: Firstly, I worked together with co-authors located in Finland, Germany, Israel, and the West Bank. Thereby, we faced some of the (cooperation and coordination) issues - partly based on cultural differences - we were investigating at the time (e.g., different perceptions of the same content of information, or disposable time). Secondly, most of the experiments discussed in the following chapters were separately or simultaneously accomplished abroad. Hence, my thesis combines cross-cultural as well as actual inter-cultural data. Thirdly, people from different disciplines and occupations supporting differing methods and research approaches contributed to the success of this interdisciplinary work. Certainly, to some extent I have also crossed my own borders while writing this thesis.

¹FUKUYAMA, F. (1995): “Trust: The Social Virtues and the Creation of Prosperity”, *The Free Press, New York*.

²KNACK, S. and P. KEEFER (1997): “Does Social Capital Have an Economic Payoff? A Cross-Country Investigation”, *Quarterly Journal of Economics*, 112, 1251-1288.

The first chapter is based on the idea to create a straightforward and symmetrical cooperation (trust) game with a continuous strategy space. First data for this study were gathered in May 2006 when we first visited the Center for the Study of Rationality at Hebrew University (Jerusalem) and Al Quds University (Abu-Dis). Later, we included additional observations from the Sichuan University (Chengdu) and the University of Helsinki (Helsinki). The chapter investigates the impact of game presentation on cooperation dependent on subject pool affiliation. Two cooperation games representing the same logical and strategic decision problem are introduced. Presented games are continuous prisoner's dilemma games where decision makers can choose an individual level of cooperation from a given range of possible actions. In the first condition, a positive transfer creates a positive externality for the opposite player. In the second condition, this externality is negative. Accomplishing an international experimental study involving subjects from Abu-Dis (West Bank), Chengdu (China), Helsinki (Finland), and Jerusalem (Israel) we test for a strategic presentation bias applying these two conditions. Subjects in Abu-Dis and Chengdu show a substantially higher cooperation level in the positive externality treatment. In Helsinki and Jerusalem no presentation effect is observed. Critically discussing our findings, we argue that comparisons across subject pools might lead to only partially meaningful and opposed results if only one treatment condition is evaluated. We therefore suggest a complementary application and consideration of different presentations of identical decision problems within research on subject pool differences.

Chapter two was planned, conducted, and written in Helsinki where I spent seven months, from September 2006 to April 2007, at the department of Psychology of the University of Helsinki. It represents a continuation of the paper presented in chapter one. Three studies predict and find that the individual's conformism values are one determinant of whether behavior is guided by other personal values or by social norms. In study one, pro-gay law reform participants are told they were either in a minority or a majority in terms of their attitude towards the law reform. Only participants who are high in conformism values conform to the group norm on public behavior intentions. In studies two and three, participants play multiple choice prisoner's dilemma games. Only participants who consider conformism values to be relatively unimportant show the expected connections between universalism values and cooperative behavior. Study three also establishes that the moderating effect of conformism values on the relation between universalism values and cooperative behavior is mediated through experienced sense of moral obligation.

Chapter three represents the work I started with at the laboratory in early 2004 when our research project on interactive decision-making involving Germans, Israelis, and Palestinians was launched. The main goal of the project is to understand how cultural differences affect

strategic behavior in a variety of strategic environments. The basic design of the study and its unique procedure and accomplishment are rooted in my diploma thesis where I developed a method to conduct inter-cultural experiments. In chapter three we experimentally investigate actions and beliefs in a trilateral trust game played among Germans, Israelis, and Palestinians. We study the influence of ethnocentrism on both trusting and reciprocity behavior. We find levels of trust, and partly of reciprocity, to differ significantly in the three national subject pools with high Palestinian and low Israeli transfer amounts. The most startling result is the fact that players' beliefs match well the actual behavior of their counterparts from their own country but that they are wrong in predicting the behavior of players located in other countries. Moreover, only slight discrimination among involved cultural groups is found. In general, the most significant differences in interactions are those between Israelis and Palestinians. Israeli senders make low transfers - to all subjects - and Israeli responders' beliefs attribute low transfers to Palestinians. In contrast, Palestinian senders make high transfers and Palestinian responders expect Israeli senders to make high transfers as well. An interaction between a Palestinian sender and an Israeli responder is likely to result in an outcome that will be perceived as positive discrimination while an interaction between an Israeli sender and a Palestinian responder is expected to create a false perception of negative discrimination. The data presented support the view that a conflict might not only be triggered by discrimination but may also be enforced by the specific social standards within different societies and ethnocentric biases in beliefs and actions.

The last chapter was accomplished in cooperation with the Department of Justice of North Rhine-Westphalia during 2004 and 2008. The main intention of this cooperation was to improve and develop the adaptiveness and certification ability of in-prison education programs dependent on reasons for ex-offenders labor market experiences. In chapter four we investigate causes and determinants of previously convicted job seekers' inferior labor market performance. The disadvantage of previously convicted applicants, an often overlooked social group, represents an important and complex issue. Despite plausibility, a proper detection of discrimination with typically available data is difficult. We experimentally investigate whether job seekers or employees holding a criminal record are less trusted and expected to be less trustworthy compared to not previously convicted workers. In addition, we compare employers' discriminative behavior against ex-offenders with their attitudes against other potential target groups of discrimination as foreigners, women, or elder workers. Our results give substantial support for a clear disadvantage of previously convicted and foreign employees who are less preferred for employment and paid significantly lower wages compared to not previously convicted (and) German workers. However, we find associated beliefs about con-

victed and foreign employees' reciprocity not to play a primary role in interactions with them. We can also show that employers' discrimination against ex-felons is mainly taste based or avoidance driven. Contrary, foreign workers are basically avoided by employers. For females and males the evidence is mixed: Employers slightly preferred females for hire and paid them higher wages. Age had practically no influence on employers' decisions. We suggest investing in prison inmates' education, its certification, as well as signalling it adequately.

Dear reader, I feel honored that you spent much of your precious time to study this work. Enjoy delving into it.

GARI WALKOWITZ

Bonn, March 29, 2010.

Chapter I

On the Prevalence of Framing Effects Across Subject Pools*

“It would be corruption to take if one could choose between ‘to take’ and ‘not to take’. It would not be a favor to give if one could choose between ‘to give’ and ‘not to give’. [...]”

MENCIUS

1 Introduction

Nowadays it is widely accepted - even by economists - that human behavior is not solely driven by the ratio of the homo economicus. Many studies have shown that subjects' behavior can be influenced amongst others by their risk attitudes, fairness or equity preferences, or by the framing of a decision problem¹. A vast body of literature demonstrates that differently *labeled* decision tasks can lead to divergent and non-rational behavior (e.g., refer to TVERSKY and KAHNEMAN, 1981; ELLIOTT, HAYWARD, and CANON, 1998; LIBERMAN, SAMUELS, and ROSS, 2004). Furthermore, other contributions have shown that subjects' behavior can be influenced even by the mere presentation of the same essential information as positive or negative (e.g., LEVIN, SCHNEIDER, and GAETH, 1998). This type of framing is commonly known as *valence framing*². In this broad field, studies dealing with public goods games creating either positive externalities (public good) or negative externalities (public bad) are well established (c.f. FLEISHMAN, 1988; ANDREONI, 1995; SONNEMANS, SCHRAM, and OFFERMAN, 1998; WILLINGER and ZIEGELMEYER, 1999; COOKSON, 2000; and PARK, 2000; DUFWENBERG, GÄCHTER, and HENNIG-SCHMIDT, 2008). Results from these publications in general suggest that experimental designs enabling positive externalities are aligned with significantly higher cooperation levels compared to setups allowing for negative externalities³.

*This chapter is based on the paper: “On the Prevalence of Framing Effects Across Subject Pools” by SEBASTIAN GOERG and GARI WALKOWITZ; re-submitted manuscript, under review.

¹See PRUITT (1967), SELTEN and BERG (1970), SELTEN (1978), KAHNEMAN and TVERSKY (1979), FEHR and SCHMIDT (1999), KONOW (2000), BOLTON and OCKENFELS (2000).

²See also ABBINK and HENNIG-SCHMIDT (2006) for a review on framing literature and framing types.

³Applying a more complex experimental design BREWER and KRAMER (1986) and SELL, CHEN, HUNTER-HOLMES, and JOHANSSON (2002) found an effect that went into the opposite direction.

In this chapter we intend to analyze subject pool affiliation as one factor leading to different levels of cooperation dependent on game presentation forms with either positive or negative externality. To maximize chances of observing behavioral differences across subject pools we conducted our study with comparable decision makers located in Abu-Dis (West Bank), Chengdu (China), Helsinki (Finland), and Jerusalem (Israel). As shown by several authors choosing subjects from different countries promises substantial cross-societal variation (see e.g., ROTH, PRASNIKAR, OKUNO-FUJIWARA, and ZAMIR, 1991; HERRMANN, THÖNI, and GÄCHTER, 2008). Our subject pools diverge according to widely used criteria developed by social scientists in order to characterize societies and, partly, in geographical distance⁴. In addition to cultural classifications, the historical and political background of Israelis and Palestinians⁵ makes them a promising testbed for investigating the link between subject pool affiliation and cooperative behavior.

Since cooperation in situations with positive or negative externality is crucial for human interaction from an individual perspective as well as from a societal point of view we will also separately compare behavior under both conditions across subject pools to evaluate the validity of findings. We will further show that our experimental approach and the awareness of the impact of subject pool differences on frame perception have importance for theoretical and practical implications. Formally identical bargaining setups might be perceived differently, evoke deviant beliefs in different subject pools, and unconsciously lead to unintended behavior. Knowing the impact of diametral frames might be essential for the design of institutions built up to moderate the relationship between involved (conflicting) parties.

So far conducted cross-societal studies typically apply experimental designs with one form of presentation. Possible, implicitly induced, presentation effects - although not in focus of the study - are not considered (e.g., ANDERSON, RODGERS and RODRIGUEZ, 2000; HENRICH, BOYD, BOWLES, CAMERER, FEHR, GINTIS, and McELREATH, 2001; BUCHAN, CROSON, and JOHNSON, 2004)⁶. To the best of our knowledge there exist only two studies taking a cross-societal perspective of framing effects into account. The first work, a questionnaire study by LEVIN, GAETH, EVANGELISTA, ALBAUM, and SCHREIBER (2001), involves Americans and Australians. Therein, American subjects stated to reduce a significantly higher amount of red meat consumption if the negative consequences of not reducing were stressed compared

⁴Contrary to geographically distant Chinese and Palestinians who are collectivistic and high context groups, Finns and Israelis - who are also located far from each other - live in more individualistic and low context societies (HALL, 1976; HOFSTEDE, 2001).

⁵At the moment, a Palestinian state does not exist. Most of our subjects are formally citizens of the states of Israel and Jordan. Nevertheless, we will refer to them as Palestinians to ease the notation.

⁶See for a specific international overview of public goods and commons dilemma studies CARDENAS and CARPENTER (2004).

to a treatment in which the positive consequences of reducing were emphasized. Contrary, Australian subjects did not respond differently to the two frames. In a second study, SELL et al. (2002) investigated the consistency and direction of framing effects across different countries. They found very similar patterns of cooperation both in the United States and in China. In both subject pools group members were more cooperative when facing resource goods dilemmas compared to a situation where they were confronted with a standard public goods give game. We will extend the approach of SELL et al. (2002) with regard to a more extensive cross-subject pool analysis and a discussion of the behavioral and methodological consequences of our findings.

For our study we conducted two series of experiments applying two continuous prisoner's dilemma games which represent different presentations of the same - strategically equivalent - decision task. That is, in both cases individuals must choose between a maximization of their own profit or to cooperate at some personal cost and increase the joint payoff. Individuals can give up an immediate benefit to sustain a resource for the other player's use. Thus, in one treatment, action creates a positive externality for the matched player. Contrary, in our second treatment, action results in a negative externality. Like a public goods dilemma our first treatment is a problem of contribution. Only with positive contributions an increase of efficiency is achieved. Similarly, our second treatment, like a commons dilemma, is a problem of consumption. The lower the share of personal consumption the higher is the efficiency. Expected utility theory suggests that these two types of presentations are equivalent since strategies and related payoffs are equivalent. However, giving and taking are psychologically different actions and findings from one set of studies may not be generalized to the other set, generally between different subject pools and especially in a cross-societal environment (BREWER and KRAMER, 1986; FLEISHMAN, 1988). This fact makes our experimental framework an appropriate tool to study presentation effects across subject pools.

Our West Bank and Chinese data show that game presentation can significantly influence decision makers' choices. In the positive externality condition substantially more cooperation is manifested in both subject pools compared to the situation with negative externality. In contrast, the experiments conducted in Helsinki and Jerusalem yielded a different result. There, on an aggregate level, no significant presentation effect can be detected. For all participant pools our data show that neither the Nash equilibrium nor the social optimal strategy is reached.

Comparing the level of cooperation under each of our two conditions across subject pools yields opposite conclusions about cooperative behavior in the different subject pools. While

subjects in Abu-Dis and Chengdu are more cooperative in the treatment with positive externality, behavior in the treatment with negative externality is more cooperative in Helsinki and Jerusalem. In contrast to this, an evaluation of all data gathered from each of the four populations shows no significant differences in cooperation levels between subject pools.

Our results shed new light on the impact of presentation conditioned by preferences and social norms in different habitats. Framing effects, in particular if they rely on different social motivations, are not robust. Therefore, we will argue that for deriving a conclusion about a subject pool's cooperative behavior, different presentations of logically identical experimental setups should be considered and evaluated adequately.

The remainder of this chapter is organized as follows: In the next part we will introduce our experimental framework consisting of two logically identical cooperation games. In the third section, we describe the procedures we applied conducting the experiments in Abu-Dis, Chengdu, Helsinki, and Jerusalem. In part four, we present subject pool-specific results. We compare data within and across populations. The final section five discusses our findings and their impact on research across different subject pools and societies.

2 Experimental framework: Two cooperation games

The two applied games are both continuous prisoner's dilemma games and public goods games in which subjects can choose an individual level of cooperation from a given range of possible actions⁷. Thus, in contrast to the classical prisoner's dilemma game, the question whether to cooperate or to defect is not a binary choice. In the first game (PDP) a player's decision creates a positive externality to the matched player's payoff, while in the second game (PDN) it induces a negative externality. In the next subsection we will describe both the PDP-game and the PDN-game in detail.

2.1 Continuous prisoner's dilemma with positive externality (PDP)

At the beginning of the game, two randomly matched players i and j obtain an integer initial endowment $X = X_i = X_j$. Each player then has the opportunity to transfer an integer part a of X , nothing, or the entire amount X to the opposite player. Both players choose $a \in \{0, \dots, X\}$ simultaneously. Each amount a , which is transferred to the paired player, will be multiplied by factor $k > 1$ yielding an efficiency gain by transferring a positive amount a . Players' payoffs consist of the initial endowment X minus the transferred amount a plus the

⁷Please refer to Appendix A for further details on the PD- and PG-nature of the two games.

obtained and k -multiplied amount a transferred by the opposite player. Formally, player i 's payoff function is given by:

$$\pi_i^{PDP} = X_i - a_i^{PDP} + k \cdot a_j^{PDP}, \text{ with } a_i^{PDP}, a_j^{PDP} \in \{0, 1, \dots, X\}, \text{ and } k > 1$$

The payoff of the opposite player j is calculated analogously. The only Nash equilibrium is $a_i^* = a_j^* = 0$. Player i anticipates player j 's choice $a_j^{PDP} = 0$ and will therefore also choose $a_i^{PDP} = 0$. The collective optimal choice is $\hat{a}_i = \hat{a}_j = X$ since it maximizes the joint payoff $\Pi^{PDP} = \pi_i + \pi_j$.

2.2 Continuous prisoner's dilemma with negative externality (PDN)

The design of the PDN-game is equivalent to the first game, but instead of choosing an amount a which is transferred *to* the opposite player, decision makers must choose an integer which is transferred *from* the other player. Again two players i and j interact simultaneously. Initially, both receive an endowment $X = X_i = X_j$. Each player then has the opportunity to transfer an integer part a , nothing, or the entire amount X *from* the matched player. Thus, again, both players simultaneously choose $a \in \{0, \dots, X\}$. The difference $X - a$, which is respectively not transferred, will be multiplied with $k > 1$. Hence, by transferring low amounts or nothing, efficiency increases. In contrast to the PDP-game, the amount a , which is transferred, is not multiplied. Players' payoffs are determined by the multiplied difference of their initial endowments X and the amount a taken by the opposite player, and the amount a which players take away from the counterpart. Formally, player i 's payoff function is given by:

$$\pi_i^{PDN} = (X_i - a_j^{PDN}) \cdot k + a_i^{PDN}, \text{ with } a_i^{PDN}, a_j^{PDN} \in \{0, 1, \dots, X\}, \text{ and } k > 1$$

Player j 's payoff is calculated analogously. The only Nash equilibrium is $a_i^* = X_j$ and $a_j^* = X_i$. Player i anticipates player j 's choice $a_j^{PDN} = X_i$ and will therefore also choose $a_i^{PDN} = X_j$. The optimal collective choice is $\hat{a}_i = \hat{a}_j = 0$ since it maximizes the joint payoff $\Pi^{PDN} = \pi_i + \pi_j$.

2.3 Equivalence of the two games

In both games player i 's payoff π_i consists of two parts - a self-determined component π_{iA} and a part π_{iB} resulting from player j 's actions. Therefore, the total payoff of player i can be stated as: $\pi_i = \pi_{iA} + \pi_{iB}$. Player i 's self-determined payoff fraction in the PDP-game is the amount $X_i^{PDP} - a_i^{PDP}$ which is not given to the other player. In the PDN-game it is the amount a_i^{PDN} that is taken away from the other player.

The foreign determined amount $k \cdot a_j^{PDP}$ for player i in the PDP-game is the amount which he receives from the matched player. In the PDN-game the foreign determined amount is the payoff fraction $k \cdot (X_i^{PDN} - a_j^{PDN})$ that the matched player leaves to him. This illustrates that in each strategy space of the two games there exists a strategy a_i or a strategy-combination $(a_i; a_j)$ that also exists in the corresponding game in terms of cooperation as well as individual and collective payoff.

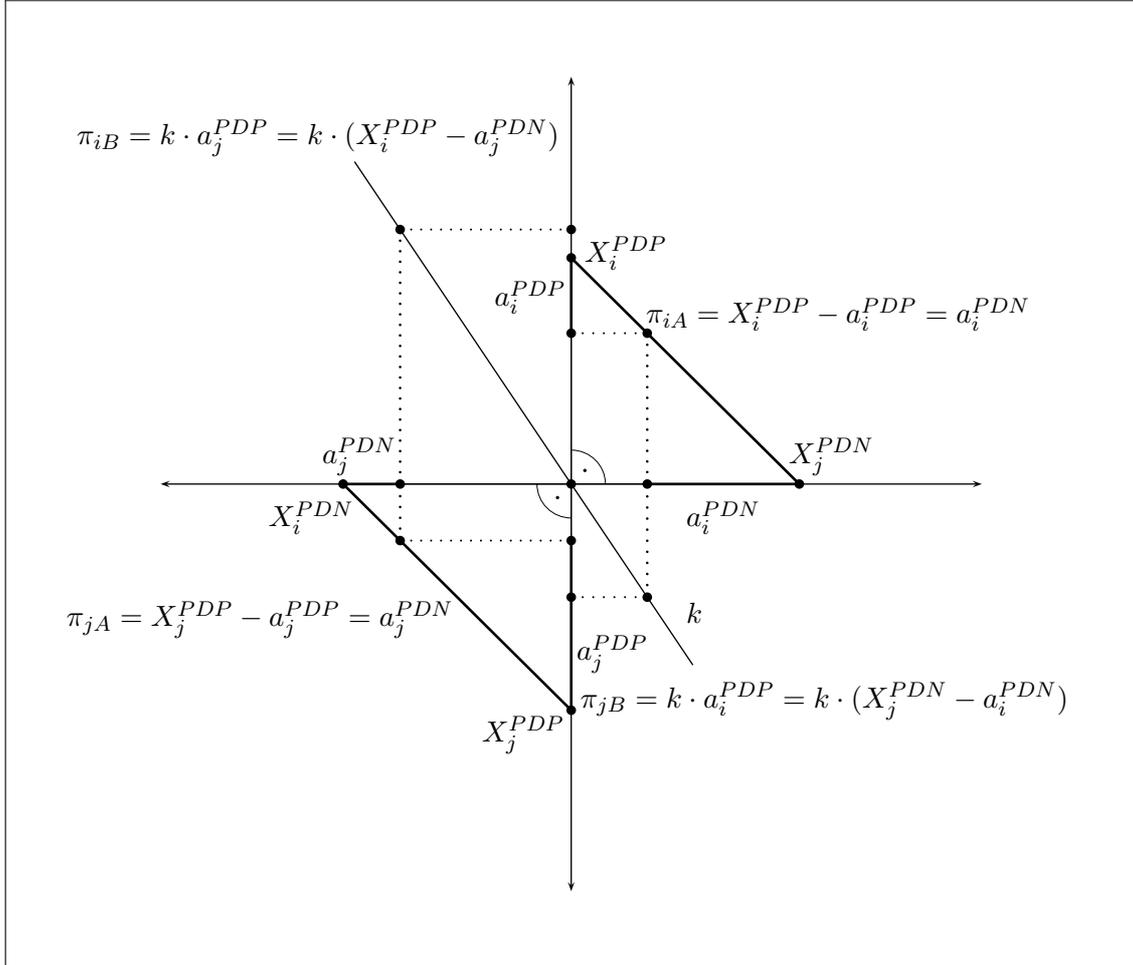


Figure 1: Graphical illustration for the equivalence of the two games.

Figure 1 displays a graphical illustration of this equivalence. We will first turn our attention to the first quadrant of the figure. The first quadrant illustrates the composition of player i 's self-determined payoff π_{iA} . In PDP π_{iA} is limited by player i 's own initial endowment X_i^{PDP} (given on the ordinate) and in PDN it is limited by the initial endowment X_j^{PDN} of the matched player j (given on the abscissa). The initial endowment X is the same in both games and for both players. Thus, X_i^{PDP} and X_j^{PDN} form an isosceles triangle as

shown in the upper right section of the figure. Player i chooses in the PDP-game the amount a_i^{PDP} (thick line on the ordinate), which is transferred to the other player. Therefore his self-determined payoff is given by $X_i^{PDP} - a_i^{PDP}$ (thin line on the ordinate). In the PDN-treatment player i can choose a_i^{PDN} (thick line on the abscissa), which ensures him the same self-determined payoff. One can see in the figure that the self-determined payoff in PDP (thin line on the ordinate) has the same size as the one in PDN (thick line on the abscissa). This is ensured by the isosceles triangle given by the initial payoffs. The third quadrant analogously illustrates player j 's self-determined payoff π_{jA} . If player i fixes his self-determined payoff as described in the first quadrant, the left over in PDN $X_j^{PDN} - a_i^{PDN}$ (thin line on the abscissa) equals the amount a_i^{PDP} transferred in PDP (thick line on the ordinate). These amounts are part of player j 's foreign-determined payoff π_{jB} and are multiplied with k which is shown in the second quadrant. The multiplier k is described as a straight line. The total amount of π_{jB} (including the multiplication with k) has to be taken from the ordinate in the second quadrant. Analogously, player i 's foreign-determined payoff is given in the fourth quadrant. This illustrates that there are one-to-one mappings f_i and f_j of player i 's or j 's strategies a_i and a_j , respectively, in the PDP-game onto the strategy spaces for i and j in the PDN-game such that a_i, a_j and $(f_i(a_i), f_j(a_j))$ yield the same payoffs in the PDP-game and the PDN-game, respectively.

3 Experimental procedures

The experiments were conducted between May 2006 and February 2007. The sessions in Abu-Dis were run at the AlQuds University located in the West Bank, close to the city of Jerusalem. We collected the Chinese data at the Sichuan University in Chengdu. Finnish data were gathered at the University of Helsinki and observations for Jerusalem were gained at the RatioLab of the Hebrew University in Jerusalem. In all universities students from different departments participated⁸.

In Abu-Dis and Jerusalem showing up for the experiment each student received a fixed payment of 25 NIS. In Chengdu [Helsinki] each subject was paid 20 YUAN [4 EURO]. At each university both games were played as one-shot games, applying the pen and paper method. We have chosen one-shot games to avoid confounding framing effects with strategic issues

⁸In Chengdu, Helsinki, and Jerusalem only subjects with very limited experimental experience were recruited (excluding previous collaborations in trust game, prisoner's dilemma, gift exchange, or public goods game experiments) for participation. Palestinian subjects had no experimental experience. The median age in Abu-Dis [Chengdu, Helsinki, Jerusalem] was 22 [23,24,25], and 55% [48%,69%,40%] females participated, respectively. We checked for possible effects of age and gender. We could not find any significant influences, neither for each subject pool separately nor for the complete sample of observations.

and to elicit behavior only conditioned to subjects' experiences and perceptions within their specific reference groups. Table 1 displays sessions, treatment conditions, and locations.

Session	Treatment	Location	Subjects
1	PDP	Abu-Dis	20
2	PDN	Abu-Dis	20
3	PDP	Jerusalem	20
4	PDN	Jerusalem	20
5	PDP	Chengdu	20
6	PDN	Chengdu	20
7	PDP	Helsinki	21
8	PDN	Helsinki	21

Table 1: Sessions, treatment conditions, locations, and subjects.

Experiments were run by local helpers comprehensively instructed and supported by the authors, who were present but stayed in the background. We are aware that this might potentially result in an experimenter effect. We decided to choose this procedure to avoid self-presentation and face-saving effects (see, e.g., BOND and HWANG, 1986) of unexperienced subjects caused by the presence of people from foreign countries. Since we are interested in the pure presentation effect this procedure seems to be justified.

Instructions were written in neutral language strictly avoiding terms like 'give', 'take', or 'cooperation'. In fact, we neither labeled, nor put disposable strategies into positive or negative light which should enhance the validity of our findings. They differed between treatments *only by the direction* of the conducted transfer. Accordingly, transfers were to be realized either *to* player j or *from* player j . This procedure ensured that only the technical presentation and not the wording or further frames could influence subjects' behavior. Dependent on the location, the instructions were either given in Arabic, Chinese, Finnish, or Hebrew⁹.

Subjects were initially endowed with $X = 10$ Talers in the opening of every game¹⁰. The multiplier k was fixed with $k = 2$. The individual payoff in the Nash equilibrium was 10 Talers, for each player. The Pareto optimum outcome generated 20 Talers, respectively. In the run of the experiment participants received no feedback on matched player's decisions.

After running the experiment two questionnaires were passed out. In the first questionnaire we asked participants for their first-order beliefs on the behavior of the matched player¹¹.

⁹To avoid translation errors regarding the task and the procedure instructions were translated by native speakers from German into the corresponding language and afterward back-translated into German by a different person (BRISLIN, 1970). For instructions see Appendix B.

¹⁰Taler=Experimental Currency. During the experiment all transfers were made in Taler. The exchange rate from Taler to NIS [YUAN, EURO] was 1 Taler = 2.5 NIS [1.5 YUAN, 0.5 EURO]. We adjusted expected hourly payoffs to the average hourly wage of a local student helper.

¹¹We are aware of the fact that stated beliefs can be biased by prior decisions already undertaken. However, since actual unbiased decisions are more valuable for this analysis, we agreed upon this procedure.

Correct beliefs were rewarded with an addition of 1 Taler. The second questionnaire covered socio-demographic questions. At the end of the session the outcome for each participant was calculated, converted into the local currency, and paid out.

4 Results

In this section we present the results of our study. Basis of our analysis is the level of cooperation exhibited by the participants in both treatment conditions. In the PDP-game it is the transferred amount to the other player (a^{PDP}) and in the PDN-game it is the amount left to the other player ($10 - a^{PDN}$). We will first take a look at the general level of cooperation found in each subject pool including decisions under both treatment conditions. Next we will investigate potential framing effects within our subject pools and then contrast actual framing effect sizes across subject pools. Afterward we will compare treatment-dependent cooperation levels across subject pools and check for consistency among treatments.

4.1 General cooperation in different subject pools

When we calculate average [expected] general cooperation levels between subject pools over *both* frames we find very similar behavior [beliefs] in all locations as 4.88 [4.40] for Abu-Dis, 4.08 [3.88] for Chengdu, 4.17 [3.62] for Helsinki, and 4.48 [3.40] for Jerusalem. No significant differences are observed for all pairwise comparisons of the aggregated data ($p \geq .22$ [$p \geq .40$], Mann-Whitney-U test, two-sided¹²). Unlike several other cross-cultural studies we do not detect significant differences between subject pools from different countries when considering all available data. In addition, we observe a positive and significant correlation between behavior and beliefs in all subject pools. The Spearman rank-correlations for each location are given by $\rho = .64$ (Abu-Dis), $\rho = .68$ (Chengdu), $\rho = .58$ (Helsinki), and $\rho = .41$ (Jerusalem) (all $p < .01$). Therefore, we conclude:

Result 1: *General levels of cooperation and their expectation do not significantly differ across subject pools. Actual behavior and associated beliefs are positively correlated.*

4.2 The impact of framing in different subject pools

Next we will explore potential framing effects in our subject pools. In Abu-Dis [Chengdu] on average 7.10 [5.50] Talers were transferred to the opposite player in PDP. Contrary, in PDN only 2.65 [2.65] Talers were left to the opponent. The observed mean differences between the two treatments are highly significant in both locations ($p < .01$ [$p < .05$], MWU). A median

¹²In the following we denote this test as ‘MWU’. All statistical tests are carried out two-sided.

analysis yields a similar result (see Figure 2 for the median levels of cooperation per treatment and subject pool). When looking at beliefs we find a comparable pattern: In both locations subjects on average expected higher cooperation in PDP (6.05 [4.55]) compared to PDN (2.75 [3.20])¹³. Given our observations in Abu-Dis and Chengdu we conclude our second result:

Result 2: *The formal presentation of the game substantially influences subjects' cooperation behavior in Abu-Dis and Chengdu. In both subject pools cooperation is significantly and economically higher under the PDP-condition than in the PDN-treatment.*

In Helsinki and Jerusalem we find a different picture. In Helsinki [Jerusalem] the shown level of cooperation is slightly higher in PDN than in PDP. In PDP on average 3.67 [4.40] Talers were transferred to the other player compared to 4.67 [4.55] which were left in PDN. However, these differences are not statistically significant ($p = .40$ [$p = .95$], MWU).

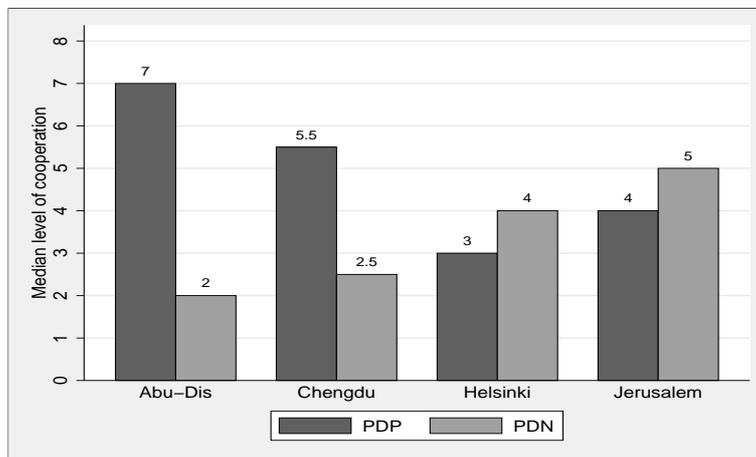


Figure 2: Median levels of cooperation per treatment and subject pool.

The same finding holds for beliefs: In Helsinki [Jerusalem] subjects on average expected similar cooperation in PDP (3.57 [3.40]) as compared to PDN (3.67 [3.40])¹⁴. Given our observations in Helsinki and Jerusalem we conclude our third result:

Result 3: *No evidence is found that the formal presentation of the game substantially influences subjects' cooperation behavior in Helsinki and Jerusalem. Both conditions imply similar levels of cooperation in both subject pools.*

Our results so far show that framing effects are not robust; in Abu-Dis and Chengdu they are detectable whereas in Helsinki and Jerusalem no evidence for a presentation bias is found. In

¹³In Abu-Dis the difference in beliefs is highly significant ($p < .01$, MWU) whereas in Chengdu mean beliefs do not differ significantly ($p = .32$, MWU).

¹⁴In both cities beliefs do not differ significantly ($p = .88$ [$p = .97$], MWU).

addition, mean differences between PDP and PDN also differ across subject pools. In order to analyze the magnitude of these differences we now compare effect sizes¹⁵ in pairs between subject pools¹⁶. Our analysis shows that the effect size in Abu-Dis [Chengdu] is significantly higher than in Helsinki (both $p < .01$) and in Jerusalem ($p < .01$ [$p = .05$]). No differences in effect sizes can be found among Abu-Dis and Chengdu as well as for a comparison between Helsinki and Jerusalem ($p = .14$ and $p = .61$). This leads to our fourth result:

Result 4: *In Abu-Dis and Chengdu subjects are significantly more sensitive to the game framing than subjects from Helsinki and Jerusalem.*

4.3 Treatment-dependent cooperation across subject pools

Based on our previous findings we will now illustrate the resulting implications for cross-subject pool comparisons. In PDP the highest cooperation level is observed in Abu-Dis (7.10), followed by Chengdu (5.50), Jerusalem (4.40), and Helsinki (3.67). Contributions in Abu-Dis are economically and statistically higher than in Jerusalem and Helsinki (both $p < .01$, MWU). Moreover, subjects in Chengdu are more cooperative than in Helsinki ($p < 0.1$, MWU) (see also Figure 2). Hence, when we only look at the positive externality frame, which is commonly used for subject pool comparisons we find clear varieties. This leads us to our fifth result:

Result 5: *In PDP, cooperation levels are higher in Abu-Dis and Chengdu compared to Helsinki and Jerusalem.*

Next, we investigate cooperation levels in PDN across subject pools. In PDN the highest mean cooperation is found in Helsinki (4.67), followed by Jerusalem (4.55), Chengdu (2.65), and Abu-Dis (2.65). Cooperation in Helsinki is significantly higher than in Chengdu and Abu-Dis ($p < .10$ and $p < .05$, both MWU) (see also Figure 2). The same holds true for Jerusalem; here cooperation is also significantly higher compared to Chengdu and Abu-Dis ($p < .05$ and $p < .10$, both MWU). Hence, we conclude:

Result 6: *In PDN, cooperation levels are higher in Helsinki and Jerusalem compared to Chengdu and Abu-Dis.*

¹⁵This is the difference in cooperation levels between PDP and PDN per location.

¹⁶We compare the differences in differences between two locations and apply a statistical test to them. To test the differences in the effect sizes we use a Monte-Carlo approximation of a two-sided permutation test with 50,000 draws. The test computes the probability for obtaining a sample with the same or a larger difference between the two effect sizes by randomly assigning each single action in PDP and PDN to one of the two locations, while at the same time keeping the condition constant. The resulting probability functions as the p -value. This procedure is analogous to ABBINK and ROCKENBACH (2006).

Our Results 5 and 6 demonstrate that observed levels of cooperation broadly depend on the game presentation form. In PDP we observe the highest level of cooperation in Abu-Dis and the lowest in Helsinki. Contrary, in PDN the order is reversed; there we find most cooperation in Helsinki and least in Abu-Dis (together with Chengdu). The differences in cooperation levels between these two locations are in both treatments significant (see Results 5 and 6). This swapping of order shows how sensitive cross-subject pool comparisons are toward game frames - results of comparisons under one game frame do not necessarily hold for other frames of the same decision task. Hence, we conclude the result section with our last finding:

***Result 7:** The elicitation and comparison of cooperative behavior across subject-pools highly depends on the chosen game presentation form. Results may adequately vary across game frames and observed subject-pools.*

5 Summary and Discussion

The aim of this work was to investigate the impact of game presentation on cooperative behavior dependent on subject pool affiliation. Merging the experimental application of two logically and strategically identical decision problems in an international setting we demonstrated that data obtained from only one presentation form might lead to only partly valid results and conclusions on subject pool-specific behavior. This finding holds especially true if results are compared and evaluated across subject pools.

Our results from Abu-Dis and Chengdu have shown that the formal presentation of a decision problem can influence subjects' choices and beliefs substantially. The cooperation level and associated beliefs are significantly higher when subjects can create positive externalities toward each other compared to a situation wherein resulting externalities are negative. In the positive condition subjects from Abu-Dis and Chengdu are more willing to transfer higher amounts to voluntarily increase mutual welfare. On average, this attitude is also expected from the opposite player. Contrary, in the negative condition subjects leave relatively less to the counterpart. In this interaction also more negative beliefs about the opponents' behavior are formed.

Experiments run in Helsinki and Jerusalem yielded different results. There, aggregated subjects' actions and beliefs appear to be externally unaffected across treatments in terms of the measured average outcome. No significant presentation effect can be verified. Finns and Israelis (even though geographically not far away located from Abu-Dis) seem to show a similar behavioral pattern in both treatments.

Our findings in Abu-Dis and Chengdu are in line with prior work on presentation effects in public good games (e.g. ANDREONI, 1995; SONNEMANS et al., 1998; WILLINGER and ZIEGELMEYER, 1999; and PARK, 2000) and with studies on goal framing (e.g., MEYEROWITZ and CHAIKEN, 1987; LEVIN, SCHNEIDER and GAETH, 1998). In these experiments the negative formulation of an identical problem has an higher impact on subjects behavior than a positive one. The observed attitudes could be connected to the concept of loss aversion and the so-called endowment effect as introduced by KAHNEMAN and TVERSKY (1979), THALER (1980), and KAHNEMAN, KNETSCH, and THALER (1990). It is possible that, even if the technical presentation of the implemented game designs was kept strictly neutral, Palestinians and Chinese perceive an amount taken away from them as a substantial loss, while they perceive an amount voluntarily given away not, or less, as a loss. In other words, Palestinians and Chinese are more sensitive to a loss induced by a second person compared to a loss induced by themselves. As a consequence of this cognition, they might react much more sensitively to the threat of a possible loss induced by the right of the second player to take away any amount as compared to the situation where they can determine themselves which amount to give away. To avoid this expected loss induced by the matched player, players take more from the matched player and thus cooperation is on a lower level in the PDN-game compared to the PDP-game. This might deliver an explanation why Palestinians and Chinese seem to obtain a higher benefit from doing a good rather than from not doing a bad deed¹⁷.

An alternative explanation for the observed behavior in Abu-Dis and Chengdu refers to the action itself. In the PDP-game action leads to cooperation whereas under the PDN-condition the opposite holds. There, action results in competitive and less efficient behavior. The difference in the sensitivity toward the given frame might stem from a different attitude toward action depending on power in general. GALINSKY, GRUENFELD, and MAGEE (2003) have shown that priming high power leads to increased action in a social dilemma regardless of whether that action had pro-social or anti-social consequences. Being primed with power incites participants to both give more to and to take more from a commonly shared resource. The different perception of own power of Palestinians and Chinese - performing notably more action in the PDP-game - and Finns and Israelis - showing a similar degree of action in both conditions - may deliver an approach to explain behavioral differences across subject pools and potentially even cultures. Future research should address this issue by linking different concepts of self image (e.g., power perception, self esteem) and situational power to decisions.

Behavior in Jerusalem and Helsinki might be rooted in the fact that subjects live in Western,

¹⁷ANDREONI (1995) argues that utility of people increases if they perceive the act of transferring as doing something good (“warm-glow”) and decreases when they perceive it as doing something bad (“cold-prickle”).

more individualistic, and low context societies. In such societies ties between individuals are loose, people are expected to look out for themselves, and behavior and beliefs are spelled out explicitly (HOFSTEDE, 2001; HALL, 1976). A series of studies has illustrated that individualists often behave more competitive and outcome-oriented in cooperation settings compared to people from collectivistic and high context subject pools (e.g., HEMESATH and POMPONIO, 1998; CARPENTER, DANIERE, and TAKAHASHI, 2004; BUCHAN, JOHNSON, and CROSON, 2006; CHUAH, HOFFMANN, JONES, and WILLIAMS, 2007; HENNIG-SCHMIDT, LI, and YANG, 2008). This finding is confirmed by our data if we only consider behavior under the positive externality condition which represents the commonly applied cross-cultural game frame. Further studies must address the cause for the similarity of behavior displayed under different presentation conditions. Do Israelis and Finns actually perceive the two games as presentation forms of the same decision problem, or do they apply different approaches leading to similar behavioral consequences and outcome?

Similar to LEVIN et al. (2001), we observe that subjects in some subject pools might respond to framing effects, while others do not. This is also reflected in associated beliefs. Hence, if different frames affect beliefs of different (culturally similar) subject pools differently and if beliefs affect behavior we might find subject-pool differences in framing effects (DUFWENBERG et al., 2006). Large cultural differences which generate well pronounced discrepancies in belief structures might therefore also enhance differences in actual behavior. In addition to this, we have shown that our findings might confront cross-societal research with new challenges: Comparing levels of cooperation under each of the conditions across subject pools might lead to opposing conclusions about society-specific behavioral attitudes. Our Palestinian and Chinese subjects display a *relatively* higher cooperation level and more positive beliefs on opponent players' contributions than involved Finns and Israelis when only the positive externalities condition is considered. Contrary, Finnish and Israeli subjects cooperate *relatively* more and state substantially higher beliefs when only the negative externalities condition is regarded. However, when all available data gathered from each of the four populations are evaluated we find no evidence that relative cooperation levels and associated beliefs differ. These striking results would not have been detected by the implementation of a mere one-sided experimental approach. Taking findings from different presentations into account might not only enrich socio-economic theory but also refine our experimental methodology.

To conclude, recognizing the impact of the presented frame might be essential for the design of culture-sensitive institutions or the conduct of international negotiations where foreign agents repeatedly interact for the first time in rapidly changing environments. Bargaining and cooperation setups might be perceived differently by decision makers depending on their

ethnic background and therefore some institutions may generate higher levels of cooperation and agreements than others¹⁸.

References

- ABBINK, K. and H. HENNIG-SCHMIDT (2006): Neutral versus loaded instructions in a bribery experiment, *Experimental Economics*, Vol. 9, 103-121.
- ABBINK, K. and B. ROCKENBACH (2006): Option Pricing by Students and Professional Traders: A Behavioural Investigation, *Managerial and Decision Economics*, 27 (6), 497-510.
- ANDERSON, L. R., Y. V. RODGERS, and R. R. RODRIGUEZ (2000): Cultural differences in attitudes toward bargaining, *Economics Letters*, Vol. 69, 45-54.
- ANDREONI, J. (1995): Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments, *Quarterly Journal of Economics*, 110(1), 1-21.
- BOLTON, G. E. and A. OCKENFELS (2000): A theory of equity, reciprocity and competition, *American Economic Review*, 90(1), 166-193.
- BOND, M. H. and K. HWANG (1986): The psychology of the Chinese people, in M. H. Bond (Ed.): *The social psychology of the Chinese people*, Oxford University Press, Hong Kong, pp. 213-266.
- BREWER, M.B., and R.M. KRAMER (1986): Choice Behavior in Social Dilemmas: Effects of Social Identity, Group Size, and Decision Framing, *Journal of Personality and Social Psychology*, 50(3), 543-549.
- BRISLIN, R. W. (1970): Back-translation for cross-cultural research, *Journal of Cross-cultural Psychology*, 1(3), 185-216.
- BUCHAN, N., R. T. A. CROSON, and E. J. JOHNSON (2004): When do fairness beliefs influence bargaining behavior: Experimental evidence from Japan and the United States, *Journal of Consumer Research*, Vol. 31, 181-190.
- BUCHAN, N.R., E.J. JOHNSON, and R.T.A. CROSON (2006): "Let's get personal: an international examination of the influence of communication, culture and social distance on other regarding preferences", *Journal of Economic Behavior and Organization*, 60(3), 373-398.
- CARDENAS, J. C. and J. CARPENTER (2004): An inter-cultural examination of cooperation in the commons, *Middlebury College Working Paper*, Middlebury College.

¹⁸For example, to maintain peace in a border region of two conflict parties both sides could be asked to take efforts to reduce their hostile armed assaults by withdrawal of armed troops or to send more unarmed peace keepers into the region. This example shows that our findings and the awareness of them have potential importance for practical application.

- CARPENTER J. P., A. DANIERE, and L. TAKAHASHI (2004): Cooperation, Trust and Social Capital in Southeast Asian Urban Slums, *Journal of Economic Behavior and Organization*, 55, 533-551.
- CHUAH, S.H., R. HOFFMANN, M. JONES, and G. WILLIAMS (2007): Do Cultures Clash? Evidence from Cross-National Ultimatum Game Experiments, *Journal of Economic Behavior and Organization*, 64(1), 35-48.
- COHEN, Y. and Y. HABERFELD (1998): Second-Generation Jewish Immigration in Israel: Have the Ethnic Gaps in Schooling and Earnings Declined? *Ethnic and Racial Studies*, Vol. 21, 507-528.
- COOKSON, S. (2000): Framing effects in public good games, *Experimental Economics*, Vol. 3, 55-79.
- DUFWENBERG, M., S.GÄCHTER, H. HENNIG-SCHMIDT (2008): The Framing of Games and the Psychology of Strategic Choice, *BGSE Discussion Paper 19/2006*, University of Bonn
- ELLIOTT, C.S., D.M. HAYWARD, and S. CANON (1998): Institutional Framing: Some Experimental Evidence, *Journal of Economic Behavior and Organization*, 35, 455-64.
- FEHR, E. and K. M. SCHMIDT (1999): A theory of fairness, competition, and cooperation, *The Quarterly Journal of Economics*, 114(3), 817-868.
- FERSHTMAN, C. and U. GNEEZY (2001): Discrimination in a segmented society: An experimental approach, *The Quarterly Journal of Economics*, 116(1), 351-377.
- FLEISHMAN, J. A. (1988): The effects of decision framing and others' behavior on cooperation in a social dilemma, *Journal of Conflict Resolution*, 32(1), 162-180.
- GALINSKY, A.D., D.H. GRUENFELD, and J.C. MAGEE (2003): From Power to Action, *Journal of Personality and Social Psychology*, 85(3), 453-466.
- HALL, E. T. (1976): *Beyond Culture*, New York: Doubleday.
- HEMESATH, M. and X. POMPONIO (1998): Cooperation and Culture: Students From China and the United States in a Prisoner's Dilemma, *Cross-Cultural Research*, 32(2), 171-184.
- HENNIG-SCHMIDT, H., Z.-Y. LI, and C. YANG (2008): Why people reject advantageous offers - Non-monotonic strategies in ultimatum bargaining: Evaluating a video experiment run in PR China, *Journal of Economic Behavior and Organization*, 65 (2), 373-384.
- HENRICH, J., R. BOYD, S. BOWLES, C. CAMERER, E. FEHR, H. GINTIS, and R. MCELREATH (2001): In search of homo economicus: Behavioral experiments in 15 small-scale societies, *American Economic Review*, 91(2), 73-78.
- HERRMANN, B., C. THÖNI, and S. GÄCHTER (2007): Antisocial Punishment Across Societies, *Science*, 319, 1362-1367.

- HOFSTEDE, G. (2001): Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations. Second Edition, Thousand Oaks, CA: Sage Publications.
- KAHNEMAN, D., J.L. KNETSCH, and R.H. THALER (1990): Experimental Tests of the Endowment Effect and the Coase Theorem, *The Journal of Political Economy*, 98(6), 1325-1348.
- KAHNEMAN, D. and A. TVERSKY (1979): Prospect theory: An analysis of decision under risk, *Econometrica*, 47(2), 263-291.
- KONOW, J. (2000): Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions, *American Economic Review*, 90(4), 1072-1091.
- LEVIN, I. P., S. L. SCHNEIDER, and G. J. GAETH (1998): All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects, *Organizational Behavior and Human Decision Processes*, 76(2), 149-188.
- LEVIN, I. P., G. J. GAETH, F. EVANGELISTA, G. ALBAUM, and J. SCHREIBER (2001): How positive and negative frames influence the decisions of persons in the United States and Australia, *Asia Pacific Journal of Marketing and Logistics*, 13(2), 64-71.
- LIBERMAN, V., S. SAMUELS, and L. ROSS (2004): The Name of the Game: Predictive Power of Reputations vs. Situational Labels in Determining Prisoner's Dilemma Game Moves, *Personality and Social Psychology Bulletin*, 30, 1175-85.
- MEYEROWITZ, B. E. and S. CHAIKEN (1987): The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, Vol. 52, 500-510.
- PARK, E. S. (2000): Warm-glow versus cold-prickle: A further experimental study of framing effects on free-riding, *Journal of Economic Behavior and Organization*, 43, 405-421.
- PRUITT, D. G. (1967): Reward structure and cooperation: the decomposed Prisoner's Dilemma game, *Journal of Personality and Social Psychology*, 7(1), 21-27.
- ROTH, A. E., V. PRASNIKAR, M. OKUNO-FUJIWARA, and S. ZAMIR (1991): Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo - An experimental study, *American Economic Review*, 81(5), 1068-1095.
- SELL, J., Z.-Y. CHEN, P. HUNTER-HOLMES, and A.C. JOHANSSON (2002): A Cross-Cultural Comparison of Public Good and Resource Good Settings, *Social Psychology Quarterly*, 65(3), 285-297.
- SELTEN, R. and C. C. BERG (1970): Drei experimentelle Oligopolserien mit kontinuierlichem Zeitverlauf, *Beiträge zur experimentellen Wirtschaftsforschung*, Vol. 2, 162-221.

- SELTEN, R. (1978): The equity principle in economic behavior, in: *Decision theory and social ethics, issues in social choice*. ed. Gottinger and Leinfellner, 289-30, Dordrecht: Reifel Publishing.
- SONNEMANS, J., A. SCHRAM, and T. OFFERMAN (1998): Public good provision and public bad prevention: The effect of framing, *Journal of Economic Behavior and Organization*, Vol. 34, 143-161.
- THALER, R.H. (1980): "Toward a Positive Theory of Consumer Choice", *Journal of Economic Behavior and Organization*, 1, 39-60.
- TVERSKY, A. and D. KAHNEMAN (1981): The framing of decisions and the psychology of choice, *Science*, Vol 211, Issue 4481, 453-458.
- WILLINGER, M., and A. ZIEGELMEYER (1999): Framing and cooperation in public good games: An experiment with an interior solution, *Economics letters*, Vol. 65, 323-328.

Appendix A: External analogy with classical PD- and PG-games

To show external analogy of our continuous games with a classical binary-choice PD we write down the 2×2 -payoff matrix form of both designs including only the extreme points of total (e.g., $a_i^{PDP} = 10$; $a_i^{PDN} = 0$) and no cooperation (e.g., $a_i^{PDP} = 0$; $a_i^{PDN} = 10$):

π_1, π_2	C₂	D₂
C₁	$k \cdot X, k \cdot X$	$0, X + k \cdot X$
D₁	$X + k \cdot X, 0$	X, X

Table 6: 2×2 -matrix, representing the prisoner's dilemma game.

The PD-condition $(1 + k) \cdot X > k \cdot X > X > 0$ is satisfied for all $k > 1$ in both games. In our experiment this condition is fulfilled, with $k = 2$. Given these parameters, by linear interpolation payoffs from the discrete payoff matrix can be obtained. Having a freely pre-determined range of possible actions a allows to obtain a non-binary measure of cooperation.

We now show external analogy of both games with a typical PG-design. The payoff function of a common 2-person PG is given by:

$$\pi_i^{PG} = X_i - a_i + k \cdot \frac{a_i + a_j}{2}, \text{ with } i \neq j, \text{ and } k > 1$$

X_i represents player i 's initial endowment. The parameter a_i is the investment into the public good. Accordingly, $X_i - a_i$ represents the investment into the private good. All investments made to the public good are multiplied by the factor k . The fraction of one half of the increased public pie is returned to both players i and j by the addition to their investments into the private good. For $k < 1$ it is rational for both players to invest nothing into the public good since the public pie shrinks. In the case of $k > 1$ both players can increase their personal income by investing into the public good. However, in this case each player has a strong incentive to free-ride hoping to reach even higher returns caused by a positive investment of the second player.

From the initial PG-equation we get:

$$\pi_i = X_i - \left(1 - \frac{k}{2}\right) \cdot a_i + k \cdot \frac{a_j}{2}$$

$$\iff \pi_i = X_i - \theta \cdot a_i + k^* \cdot \theta \cdot a_j, \text{ with } \theta = 1 - \frac{k}{2}, \text{ and } k^* = \frac{k}{2 \cdot \left(1 - \frac{k}{2}\right)}$$

The payoff-function of the PDP-game was given in equation by:

$$\pi_i^{PDP} = X_i - a_i^{PDP} + k \cdot a_j^{PDP}$$

It is evident that both games are of the same type: A PG-game with parameter k^* is formally similar to the PDP-game with parameter k . Because of internal equivalence among PDP and PDN it is obvious that the PDN-game is a PG too. Contrary to the PG-game, in PDP and PDN there is no back flow of own investments. Thus, each $a_i > 0$ is transferred directly to the opposite player thereby providing a lower individual incentive to cooperate.

Appendix B: Instructions for the experiment (for PDP and PDN)

Introduction

Thank you for taking part in this experiment. Please read these instructions very carefully. It is very important that you do not talk to other participants for the time of the entire experiment. In case you do not understand some parts of the experiment, please read through these instructions again. If you have further questions after this, please give us a sign by raising your hand out of your cubicle. We will then approach you in order to answer your questions personally.

To guarantee your anonymity you will draw a personal code before the experiment starts. Please write this code on top of every sheet you use during this experiment. You will later receive your payment from this experiment by showing your personal code. This method ensures that we are not able to link your answers and decisions to you personally.

During this experiment you can earn money. The currency within the experiment is 'Taler'. The exchange rate from Taler to CURRENCY is:

$$\boxed{1 \text{ Taler} = \text{XX CURRENCY}}$$

Your personal income from the experiment depends on both **your own** decisions and on the decisions of **other** participants. Your personal income will be paid to you in cash as soon as the experiment is over.

During the course of the experiment, you will interact with a randomly assigned other participant. The assigned participant makes his/her decisions at the same point in time as you do. You will get no information on who this person actually is, neither during the experiment, nor at some point after the experiment. Similarly, the other participant will not be given any information about your identity. You will receive information about the assigned participant's decision after the entire experiment has ended.

After the experiment, please complete a short questionnaire, which we need for the statistical analysis of the experimental data.

Description of the experiment (PDP)

In this experiment you are randomly matched with another participant. You act as **Person A**, and the randomly assigned other participant acts as **Person B**. You and Person B must simultaneously make a similarly structured decision.

Person A and Person B first receive an initial endowment of **10 Talers**.

You now have the opportunity to transfer any part of your endowment to Person B. You can only transfer integer amounts - thus, you can only choose amounts $a_A \in [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$.

The amount you transfer to Person B is **doubled**. That means that Person B receives **twice the amount you have transferred** to him/her.

The randomly assigned participant acting as Person B is given exactly the same alternatives as you have. He/she also has the possibility to transfer any amount to you. The amount Person B transfers to you is also doubled. That means that **you receive twice the amount Person B has transferred** to you.

You will make your decisions **simultaneously**. During the course of the experiment neither person receives any information concerning the decision of the other person.

How the income is calculated

Your personal income can be calculated as follows:

Initial endowment
- amount you choose to transfer to Person B
+ twice the amount b Person B transferred to you
= your personal income

Description of the experiment (PDN)

In this experiment you are randomly matched with another participant. You act as **Person A**, and the randomly assigned other participant acts as **Person B**. You and Person B must simultaneously make a similarly structured decision.

Person A and Person B first receive an initial endowment of **10 Talers**.

You now have the opportunity to transfer any part of Person B's endowment to yourself. You can only transfer integer amounts - thus, you can only choose amounts $a_A \in [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$.

The remaining amount - that is the amount that you do not transfer from Person B's endowment to yourself - is **doubled**. This means that Person B receives **twice the amount that you do not transfer** from him/her.

The randomly assigned participant acting as person B is given exactly the same alternatives as you have. He/she also has the possibility to transfer any amount to himself/herself. The remaining amount that he/she does not transfer from your endowment to himself/herself is **doubled**. This means that you receive **twice the amount that he/she does not transfer** from you.

You will make your decisions **simultaneously**. During the course of the experiment, neither person receives any information concerning the decision of the other person.

How the income is calculated

Your personal income can be calculated as follows:

+ amount you choose to transfer from Person B to yourself
+ twice the amount Person B did not transfer from your endowment to himself/herself
= your personal income

Chapter II

The Moderating Effect of Conformism Values on the Relations between Other Personal Values, Social Norms, Moral Obligation, and Individual Cooperative Behavior*

“The value concept... [is] able to unify the apparently diverse interests of all the sciences concerned with human behavior.”

ROCKEACH (1973)

1 Introduction

Research on the relation between internal dispositions and external behavior has confirmed that socially desirable or normative behaviors are likely to be performed whether an actor possesses the corresponding dispositions or not (e.g., JONES and DAVIS, 1965). In the domain of values research, BARDI and SCHWARTZ (2003) suggest that the more popular or normative a behavior, the weaker its connection to personal values. The reason for this is that normative behaviors are more likely to be guided by social forces than by the individual’s personal values. However, recent research has added the important qualification that, at least in the domain of values research, the phenomenon is true primarily of people high in Conformism values (LÖNNQVIST, LEIKAS, PAUNONEN, NISSINEN, and VERKASALO, 2006).

In the introductory sections that follow, we first present a particular theory of values and previous results connecting values to behavior. We then present research on how social norms can guide behavior. Finally, we explain why Conformism values can determine, in part, whether individual behavior is governed by other personal values or by social norms. The three principal goals of our research are to (a) show that Conformism values moderate the influence of social norms on behavioral intentions, (b) show that Conformism values moderate the influence of Universalism values on individual cooperative behavior, and (c) explore the

*This chapter is based on the paper: “The Moderating Effect of Conformism Values on the Relations between Other Personal Values, Social Norms, Moral Obligation, and Single Altruistic Behaviors” by JAN-ERIK LÖNNQVIST, GARI WALKOWITZ, PHILIPP WICHARDT, MARJAANA LINDEMAN, and MARKKU VERKASALO; forthcoming in *British Journal of Social Psychology*.

process through which Conformism values moderate the influence of Universalism values on individual cooperative behavior.

Values and Behavior

Human values are typically defined as trans-situational goals, varying in importance, that serve as guiding principles in social interactions. Summarizing past research on values, SCHWARTZ (1992) defines values as concepts or beliefs that act as standards of what is most desirable when evaluating events, behaviors, and persons. In contrast to attitudes, that represent an individual's degree of like or dislike for an item, values transcend specific situations, are ordered in a person in a hierarchy of importance, set standards of desirability, and are less numerous and more central to personality than attitudes are.

SCHWARTZ (1992, 2005a) and SCHWARTZ and SAGIV (1995) have identified 10 motivationally distinct basic values that are recognized in over 50 languages and 70 countries. These 10 basic values are Benevolence, Tradition, Conformism, Security, Power, Achievement, Hedonism, Stimulation, Self-Direction, and Universalism¹. The social and psychological conflict and congruity people experience when pursuing values give rise to a somewhat predictable structure of relations among the 10 values (SCHWARTZ, 1992, 2005b). This means that actions that express a particular value may have consequences which are compatible or incompatible with other values. For instance, actions that promote Stimulation values tend also to contribute to freedom (Self-Direction), whereas they tend to be in conflict with satisfying Tradition values.

The definition of values suggests that values, as standards of the desirable, could be expected to guide behavior. However, there is some dispute regarding the role of values in guiding behavior (for a review, MAIO, OLSON, BERNARD, and LUKE, 2003), with some theorists positing the importance of values in that regard (ROKEACH, 1973; SCHWARTZ, 1992), and others denying it (e.g., KRISTIENSEN and HOTTE, 1996; MCCLELLAND, 1985). The processes through which values guide behavior (e.g., when a specific value or a set of values is activated) are also not clear. Although some theorists have held that values guide behavior primarily through attitudes (e.g., AJZEN, 1991; KRISTIENSEN and HOTTE, 1996), some empirical results suggest that values may also guide behavior independently of attitudes (MAIO and OLSON, 1995; VERPLANKEN and HOLLAND, 2002); in particular, some values may influence behavior independent of attitudes as values are thought to give rise to a sense of moral obligation to perform the behavior so as to preserve one's sense of self-worth (e.g., BECK and AJZEN, 1991; MAIO and OLSON, 1995; SCHWARTZ, 1977; SCHWARTZ and TESSLER, 1972).

¹See Appendix for a short description of each value type.

BARDI and SCHWARTZ (2003) connect values to informants' reports of behaviors. In that study, some of Schwartz' ten value types strongly predict behavior, but others do so only marginally. For instance, Stimulation and Tradition values are strongly related to peer and spouse reports of stimulation seeking and traditional behaviors, respectively. In contrast, Benevolence values are only marginally related to benevolent behaviors. The authors argue that these differences in value-behavior relations, in part, stem from normative pressure to perform certain actions.

Normative Influence on Behavior

Personal values may sometimes not guide behavior because behavior is often determined by social forces. People seldom make their decisions in isolation, but often look at others to guide their thoughts and actions (e.g., AJZEN, 1991; ASCH, 1952; SHERIF and SHERIF, 1964; TURNER, 1991). There are two basic reasons for why a person might yield to group pressure (see CIALDINI, RENO, and KALLGREN, 1990). One reason is informational influence: people may yield to others because they trust the judgments of others more than they trust their own judgments. This type of informational influence is a way of defining a position in the face of limited information, and may, if internalized by the individual, lead to genuine attitude change.

The other reason is normative influence: people may be influenced by norms because they want to fit in with the majority. Normative influence does not imply genuine attitude change; that is, the social norm is not internalized as a morally obliging personal norm (SCHWARTZ, 1977), but rather a strategic effort on behalf of an individual to be accepted and to avoid social censure. The assumption that a minority position is aversive (can lead to disapproval, hostility, rejection) explains normative influence. To avoid social punishment, people might conform to the majority position in public, regardless of private opinions. Normative influence is, therefore, assumed to cause changes in public but not private behavior. Informational influence, on the other hand, is assumed to also cause changes in private behavior. In previous empirical research, people have been found to conform to the in-group members' responses in both private and public settings, although the effects have been stronger for public settings (e.g., ABRAMS, WETHERELL, COCHRANE, HOGG, and TURNER, 1990; INSKO, DRENAN, SOLOMON, SMITH, and WADE, 1983).

The Moderating Effects of Conformism values

According to BARDI and SCHWARTZ (2003), social norms can exert strong external pressure that obscures the relation between personal values and behavior. Behaviors with a moral dimension are particularly pertinent in this regard because they are prone to be influenced by

strong social norms. For example, people who are otherwise low in Universalism values might feel compelled by others to help the needy (SCHWARTZ, 1977). Similarly, the social context may induce those who are normally high in Universalism to behave in an incongruent (i.e., less cooperative) manner (see DARLEY and BATSON, 1973; BERNARD, MAIO, and OLSON, 2003; MAIO, OLSON, ALLEN, and BERNARD, 2001).

LÖNNQVIST et al. (2006) present data consistent with the suggestion that social norms can obscure the true relation between personal values and cooperative behavior, but with the important qualification that the phenomenon is true primarily of people high in Conformism values. They show that personal values predict peer-evaluations of cooperative behavior, but only in the group of people who consider Conformism values to be relatively unimportant. Furthermore, they show that Conformism values moderate whether or not personal self-transcendence (Universalism and Benevolence) values influence the experience of regret over decisions engendering adverse consequences in hypothetical scenarios. Only if Conformism values are low, personal self-transcendence values predict regret. These results are interpreted as suggesting that only those low in Conformism values base their evaluations on their personal self-transcendence values.

Purpose of the Present Research

Our principal goal in this research is to extend research on Conformism values as moderator between other personal values and behavior. One reason other personal values do not influence the behavior of those high in Conformism values could be that those people are more prone to act in congruence with perceived social norms. This means that those high in Conformism values should be more sensitive to a manipulation of social norms; that is, more willing to conform to the group norm in their display of behaviors. Although this has been suggested or assumed in previous research (BARDI and SCHWARTZ, 2003; LÖNNQVIST et al., 2006), the idea has, to the best of our knowledge, not been empirically tested directly. In Study 1, we expect those high in Conformism values to move toward the group norm in their behavior intentions. We do not expect the group norm to influence people low in Conformism values. Furthermore, Study 1 tests the idea that Conformism values may involve a strategic effort to be accepted (greater susceptibility to normative influence), but not more to trust in the judgment of others (greater susceptibility to informational influence). If this is the case, then Conformism values should moderate the influence of group norm on public, but not private, behavior intentions.

In Study 2, we seek to show that Conformism values moderate the relation between Universalism values and individual cooperative behavior in specific situations. Based on the work of

SCHWARTZ (1996), we expect Universalism values to correlate with cooperative behavior in a prisoner's dilemma game. However, based on the work of LÖNNQVIST et al. (2006), we expect this to be true only in the group of people low in Conformism values; concerning people high in Conformism values, we do not expect Universalism values to predict cooperative behavior. Universalism values should, in our experimental design, be the most important value predicting cooperative behavior, as the defining goals of Universalism is the welfare of all people, and the cooperative behavior is directed towards an anonymous stranger. By contrast, the other self-transcendence value, Benevolence, emphasizes the welfare of close others (family, friends), and is thus less relevant for the present design.

In view of earlier research, we want to emphasize one very important difference between the study by LÖNNQVIST et al. (2006) and the present Study 2 (and Study 3). In the former study, peer ratings of benevolence, honesty, and aggressiveness are used as an approximation of cooperative behavior. This is somewhat disconcerting, as values, due to their nature as explicit motives, have previously been argued to predict immediate specific responses to specific situations, but not spontaneous behavioral trends over time (MCCLELLAND, KOESTNER, and WEINBERGER, 1989). Therefore, it is important to show that Conformism values moderate the influence of other values on actual choice behavior; that is, the type of behavior that values are usually thought to influence. Furthermore, ratings of traits are influenced by many sources of variance (e.g., GUIMOND et al., 2007; ROSS and NISBETT, 1991), and are usually not very good predictors of single behavior in specific situations (e.g., JACKSON and PAUNONEN, 1985; MISCHEL and PEAKE, 1982; MURRAY, 1938). Therefore, strictly speaking, the study by LÖNNQVIST et al. (2006) only shows that Conformism values moderate the influence of other values on trait ascriptions made by peers, but not that Conformism values moderate the influence of other values on actual cooperative behavior.

A further contribution of Study 2 is that we look at two separate types of cooperative behavior: in one prisoner's dilemma game cooperative behavior requires that one refrains from anti-social action (taking from the other person), whereas in the other game cooperative behavior requires pro-social action (giving to the other person). It seems possible that the moderating effect of Conformism values would depend on the type of cooperative behavior involved.

In Study 3 we expand upon Study 2 by looking at the process through which Conformism and Universalism values influence behavior in a prisoner's dilemma game. In order to do this, we introduce experienced sense of moral obligation as a mediator variable between Universalism values and behavior. As noted above, values are thought to give rise to a sense of moral obligation to perform the behavior in order to preserve one's sense of self-worth

(e.g., BECK and AJZEN, 1991; MAIO and OLSON, 1995; SCHWARTZ, 1977; SCHWARTZ and TESSLER, 1972). We hypothesize that the Universalism values give rise to such a feeling of moral obligation, but with the important qualification that the phenomenon is true primarily of people low in Conformism values. Universalism values are not expected to give rise to a feeling of moral obligation among those high in Conformism values.

2 Study 1: Conformism Values and Social Norms

In this study we evaluate the moderating effects of Conformism values on the ability of social norms to predict behavior intentions. The methods of Study 1 are adapted from a study by HORNSEY, MAJKUT, TERRY, and MCKIMMIE (2003). That is, students who support pro-gay law reform are given the information that, on the whole, students from their university either agree or disagree with their attitudinal position. After this, participants complete a questionnaire assessing their willingness to engage in private and public behaviors consistent with their original attitude. We expect that people high in Conformism values will adjust their public behavior intentions to be consistent with group norms. Concerning people low in Conformism values, we do not expect the group norm manipulation to influence their behavioral intentions.

2.1 Method

Participants

The initial sample in this study consisted of 73 students attending an introductory course in research methods in psychology at the Open University of Helsinki. Participants indicated on a scale from 1 (strongly oppose) to 8 (strongly support) their position on the eligibility of gay couples for assisted reproductive technology (there was a lot of discussion of such a law reform in Finland in 2006). Following HORNSEY et al. (2003), only those participants who identified themselves as being pro-gay law reform (scoring 5 or higher) were included in the final sample, because the behaviors we sought to predict (e.g., vote for a party supporting law reform) would not make sense for those opposed to the law reform. The final sample thus consisted of 50 respondents (49 females), with a mean age of 28.9 years ($SD = 9.9$). The participants were administered the questionnaires during class time.

Group norm was manipulated after participants had completed the pre-manipulation measures as well as the values questionnaire. This was done by presenting a diagram to participants that ostensibly summarized the attitudes of University of Helsinki students toward the pro-gay law reform. Students were portrayed to be either strongly in favor of gay law reform

or strongly opposed to gay law reform. In the group support condition, 67% were in favor of law reform, 19% opposed, and 14% were undecided. In the group opposition condition, 67% opposed, 19% supported, and 14% were undecided.

Measures

Pre-manipulation measures. Besides indicating their attitude toward pro-gay law reform, participants also indicated the perceived societal support of their attitude. Following HORNSEY et al. (2003), participants indicated on a scale from 1 (*not at all*) to 8 (*very much*) to what extent they felt that society agreed with their opinion.

Values. Values are measured with the Portraits Values Questionnaire (SCHWARTZ et al., 2001). This 40-item questionnaire consists of descriptions of a hypothetical person in terms of his or her goals and aspirations. Respondents are to rate their similarity to the person on each item, using a 6-point scale ranging from 1 (not like me at all) to 6 (very much like me). Proportional sum variables are used to correct for individual differences in scale use (VERKASALO, TUOMIVAARA, and LINDEMAN, 1996). The values scores, therefore, represent the relative importance to the person of each value type compared to the other value types, with the average score for all 10 value scales being 1.00. The four-item Conformism scale ($M = 0.89$, $SD = 0.22$) has a good reliability of .69 (in comparison, in validating the Portraits Values Questionnaire, SCHWARTZ et al. (2001) report a mean reliability of .48 for the Conformism scale).

Behavior intentions. After the manipulation of group norm, participants completed six items assessing their willingness to act out their attitudes in the private and public domain. Participants were asked to indicate on a scale from 1 (*not at all willing*) to 8 (*very willing*) how willing they would be to perform a number of activities provided they had time and the opportunity to do so. Following HORNSEY et al. (2003), the three private behaviors were signing a petition, voting in a referendum, and voting for a political party that had pro-gay law reform policies. The first factor extracted in principal components factor analyses of the three private behavior intentions items accounted for 74% of the variance of these three items. The three items loaded between .78 and .86 on this factor, and the factor score was used as measure of private behavior intentions. The three public behaviors were signing a letter to the editor, distributing information leaflets, and attending a rally in favor of gay law reform. The first factor extracted in principal components factor analyses of the three public behavior intentions items accounted for 76% of the variance of these three items. The three items loaded between .73 and .95 on this factor, and the factor score was used as measure of public behavior intentions.

2.2 Results and Discussion

Group norm is effects-coded, such that group opposition is coded -1 and group support is coded 1. Following HORNSEY et al. (2003), in order to rule out the effects of attitude strength and the pre-manipulation perception of societal support, we regress public and private behavior intentions on those variables and save the residuals as our new indicators of public and private behavior intentions. This strategy removed any individual differences due to attitude strength or pre-manipulation perception of societal support (see PAULHUS and JOHN, 1998).

Two sets of hierarchical multiple regression analyses are performed to examine the moderating effects of Conformism values on the effects of the group norm manipulation in predicting behavior intentions. In all analyses, continuous variables were centered to facilitate interpretation of the interaction results (AIKEN and WEST, 1991). Group norm and Conformism values, are entered in Step 1 of the regression analysis and their interaction is entered in Step 2. In the prediction of private behavior intentions, the group norm (support vs. opposition) is marginally statistically significant, $\beta = .27$, $t(3,46) = 1.91$, $p = .07$. Participants in the group support condition were more intent to privately act in favor of the pro-gay law reform than participants in the group opposition condition.

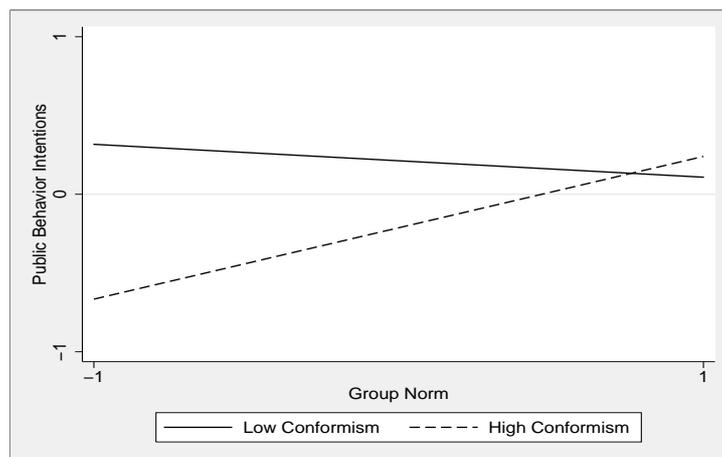


Figure 1: Public behavior intentions regressed on the group opposition (= - 1) vs. group support (= + 1) manipulation for low and high Conformism groups (Study 1). Low score = 1 SD below the mean; high score = 1 SD above the mean.

As for public behavior intentions, neither group norm ($\beta = .18$, n.s.) nor Conformism values ($\beta = -.22$, n.s.) are statistically significant predictors of public behavior intentions. However, the expected interaction between group norm and Conformism values, entered in Step 2 of the regression analysis, is marginally statistically significant ($\beta = .28$, adjusted R^2 increased from

.00 to .06, the F for the model changed from 1.10 to 2.02, F for the change in $R^2 = 3.73, p = .06$). We select data points for estimating regression lines at $\pm 1 SD$ for predictors of the equation (AIKEN and WEST, 1991). Tests of simple slopes reveal that the association between group norms and behavior is strong when Conformism values are high, $B = 0.40, SE = 0.19, \beta = .41, t(3,46) = 2.08, p < .05$, but disappear when Conformism values are low, $B = -.013, SE = 0.19, \beta = -.14, t(3,46) = -0.69, p = .49$. Figure 1 provides simple regression lines of public behavior intentions as a function of group norm at high and low ends of Conformism values. The figure shows that people high in Conformism values are reluctant to act in publicity in favor of gay law reform when they are told that the majority is against the law reform. However, when people high in Conformism values are told that the majority supports the law reform, they are as eager to act publicly in favor of the reform as those low in Conformism values are.

Our findings indicate that the group norm affects public behavior intentions, but only in the group of people who are high in Conformism values. These results thus directly support the view that social norms can influence people high in Conformism values to a degree that obscures the true influence of attitudes (recall that all participants were pro-gay law reform) on behavior. People high in Conformism values adjust their public behavior intentions to be congruent with the group norm. In contrast, no such effect is found for those low in Conformism or for private behavior intentions.

3 Study 2: Conformism Values and Cooperative Behavior

In Study 2, we evaluate the moderating effects of Conformism values on the relation between Universalism values and cooperative behavior. As explained in the introduction, we expect Universalism values to correlate with cooperative behavior in a prisoner's dilemma (PD) game (SCHWARTZ, 1996). Universalism values share the same motivational goal as Benevolence values, but Universalism values are distinguished from Benevolence values in the fact that Universalism values are concerned with the welfare of all people, whereas Benevolence values are concerned only with the welfare of those with whom one is in frequent personal contact. Therefore, in a situation such as a one-shot PD-game, where one interacts only once with a stranger whose identity one will never know, only Universalism values would be expected to predict cooperative behavior. However, based on the work of LÖNNQVIST et al. (2006), we expect Universalism values to predict cooperative behavior only in the group of people low in Conformism values.

3.1 Method

Experimental Design and Procedure

In Study 2 participants sequentially played two one-shot prisoner's dilemma games with multiple choice. One game was the 'Prisoner's Dilemma Give' (PDG) game and the other the 'Prisoner's Dilemma Take' (PDT) game (see GOERG and WALKOWITZ, 2008). Both games are two-person variants of the standard Public Goods Game and allow participants to simultaneously choose a cooperation level from a range of disposable strategies. In both games, participants are anonymously and randomly matched with a second player. Half of the participants played the PDG first and the other half the PDT first.

The currency applied in the experiment was "Taler" (1 Taler = 0.50 Euro). In PDG participants received an initial endowment of 10 Talers. They were then given the opportunity to transfer as many Talers (0 - 10; natural numbers) as they wished to the other participant. The amount the participants transferred was doubled by the experimenters. In PDT, both participants were again given 10 Talers and participants were allowed to arbitrarily transfer Talers (0 - 10; natural numbers) from that other participant to themselves. In this game, the amount that participants did not transfer from the other participant was doubled by the experimenters. Although they were presented differently, the two games are strategically identical, as they involve the same number of players, have the same action space and payoff matrix (see GOERG and WALKOWITZ, 2008, for more detailed information). Although they were presented differently, the two games are strategically identical, as they have the same number of players, action space, information structure, payoffs, Nash-equilibrium, and Pareto Optimal outcome (see GOERG and WALKOWITZ, 2008, for more detailed information). The amount transferred in PDG and the amount not transferred in PDT constituted our measures of cooperative behavior. It is important to note that although participants were paired, this happened only after the participants had made all their decisions. Thus, each participant constituted an independent observation.

In the first session, participants first played the PDG game and afterward the PDT game. The order of the games was reversed in the second session. When playing the first game, participants did not know about the upcoming second game. All participants were given exactly the same tasks and participants were told that the randomly assigned other participant was given exactly the same alternatives as they were given themselves. Participants were randomly assigned a different partner for each game. At no stage, the identity of the randomly assigned other participant was revealed. Moreover, participants did not get to know their partners' decisions until they were paid at the end of the session.

Participants

Participants were students from various faculties (e.g., social sciences, arts, natural sciences) at the University of Helsinki. Participants were contacted through university mailing lists and notices put up on general university notice boards. The notice for the study presented the study as an opportunity to earn money by taking part in a decision experiment (no hints as to the nature of the game were given). Furthermore, every participant was promised (and given) a four Euro show-up fee. Participants registered for the experiment by e-mail and were asked to show up at the given date and time. 21 participants took part in each session. However, one participant skipped a page of the values questionnaire and was omitted from further analyses. This gave us a total of 41 participants (28 female), with a mean age of 24.8 years ($SD = 4.5$ years).

On showing up, participants were randomly seated in a large classroom. The desks were organized in a way that prevented participants from looking at each other while completing the tasks (e.g., seated facing the walls). Participants first played the two different variants of the PD-games. Instructions for the two PD-games were both dealt out at each stage to participants on paper and read out aloud. After having played both PD-games participants were administered a values questionnaire and a demographic questionnaire. After having completed the questionnaires, participants left one at a time and were paid off outside the classroom (show up fee plus earnings in the decision experiment).

Measures

Values are measured with the Portraits Values Questionnaire (SCHWARTZ, MELECH, LEHMANN, BURGESS, HARRIS, and OWENS, 2001). To correct for individual differences in scale use, proportional sum variables are used (VERKASALO, TUOMIVAARA, and LINDEMAN, 1996). The descriptives of the 10 values types are given in Table 1. Although some of the reliabilities are quite low, Conformism and Universalism values, to which our hypotheses pertain, have good reliabilities (.67 and .86, respectively).

3.2 Results and Discussion

First, we explore whether the type of game (treatment condition) or the order of the games affect the transfer decisions. In order to do this, we conduct a repeated measures ANOVA with the game (PDG vs. PDT) as within-participants variable and the order as between-participants variable (in this and subsequent analyses, the amount transferred in the PDG is compared with the amount not transferred in the PDT). Neither the treatment ($F = 0, p = 1$), nor the order ($F = 1.21, p = .28$), nor their interaction ($F = 0.47, p = .83$) is statistically

significant. A non-parametric statistical analysis yields similar results: Neither the treatment ($p = .407$, Mann-Whitney-U-test, two-sided), nor the order ($p = .774$, Wilcoxon Signed Ranks Test, two-sided), nor their interaction ($p = .826$, Wilcoxon Signed Ranks Test, two-sided) is statistically significant. Next, we enter Conformism and Universalism values as covariates, and add all two-way and three-way interactions as between-participants variables. However, none of the between-participants variables shows a statistically significant interaction with treatment (no within-participants effects are statistically significant, all $F < 2.17$, n.s.). Thus, the predictive power of Conformism, Universalism, order, and all their two-way and three-way interactions do not depend on the treatment. Transfer decisions are therefore averaged over treatment and order. The mean transferred amount (across treatments and order) is 4.1 Taler ($SD = 3.0$).

The correlations between the ten value types and cooperative behavior are shown in Table 1. As expected, Universalism values positively correlate with the transferred amount ($r = .34, p < .05$). Although the correlation between Self-Direction values and cooperative behavior is marginally significant ($r = -.28, p = .07$), this correlation should be interpreted with caution considering the very low reliability (Cronbach's $\alpha = .21$) of the Self-Direction values scale.

Value Types	<i>M</i>	<i>SD</i>	α	<i>r</i>
Universalism	1.22	0.19	.86	.34*
Benevolence	1.11	0.17	.70	.20
Conformism	0.81	0.23	.67	-.09
Tradition	0.70	0.20	.36	.10
Security	0.99	0.20	.65	-.22
Power	0.73	0.19	.57	-.08
Achievement	0.94	0.25	.82	-.09
Hedonism	1.02	0.25	.93	.05
Stimulation	1.02	0.23	.91	-.10
Self-Direction	1.19	0.17	.21	-.28

Table 1: Descriptive statistics for the 10 value scales and their correlations with cooperation behavior in two multiple choice Prisoner's Dilemma games (Study 1), * $p < .05$.

Recall that Conformism might be expected to affect whether one's decisions are influenced by other personal values and, therefore, to moderate the relation between other personal values and cooperative behavior. As we have reasoned earlier, we expect correlations between Universalism values and cooperative behavior to be particularly strong in low Conformism participants. Hierarchical multiple regression analyses are performed to examine the moderating effects of Conformism values on Universalism values in the prediction of cooperative behavior (see AIKEN and WEST, 1991). Of the variables entered in Step 1, Universalism

values strongly predict ($\beta = .60, p < .01$) cooperative behavior, but Conformism values do not ($\beta = .04, n.s.$). However, the predictive power of Universalism values is qualified by the interaction between Universalism and Conformism values. The interaction between Universalism and Conformism values, entered in Step 2 of the regression analysis, is marginally statistically significant ($\beta = -.38$, adjusted R^2 increased from .08 to .15, the F for the model changed from 2.86 to 3.33, F for the change in $R^2 = 3.83, p = .06$). Simple slope analysis is conducted to illustrate the nature of the interaction (see AIKEN and WEST, 1991). Tests of simple slopes reveal that the association between Universalism and cooperative behavior is stronger when Conformism values are low, $B = 12.16, SE = 3.97, \beta = .79, t(3,37) = 3.06, p < .01$, relative to when Conformism values were high, $B = 6.13, SE = 2.68, \beta = .40, t(3,37) = 2.29, p < .05$, suggesting that universalism values guide cooperative behavior more strongly to the extent that Conformism values are low.

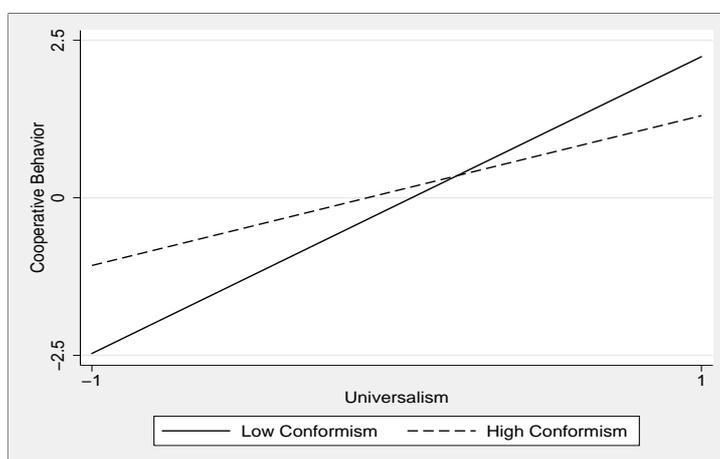


Figure 2: Cooperative behavior in two Prisoner's Dilemma games regressed on low and high Universalism scores for low and high Conformism groups (Study 2).
Low score = 1 *SD* below the mean; high score = 1 *SD* above the mean.

Figure 2 shows that cooperative behavior increases as a function of Universalism values, but more strongly for those low in Conformism; the level of Universalism has far less impact on the decision of those who consider Conformism values to be important. Conformism values do not moderate the influence of any other values on cooperative behavior.

4 Study 3: Mediation of Moderating Effects

The results of Study 2 are consistent with those of LÖNNQVIST et al. (2006), indicating that Conformism values moderate the influence of Universalism values on cooperative behavior. In Study 3 we seek to explore the process by which this happens. Values have been shown

to influence behavior, in part, by giving rise to a sense of moral obligation (e.g., MAIO and OLSON, 1995). SCHWARTZ (1977) has shown that personal values give rise to feelings of moral obligation to act in a way consistent with one's personal values, so as to enhance or preserve one's sense of self-worth. However, those high in Conformism values have been shown neither to act in a way that is consistent with their other personal values (Study 2), nor to anticipate much self-blame or regret over decisions inconsistent with their values (LÖNNQVIST et al., 2006). These results might imply that the other personal values of those high in Conformism are not enforced by an experienced sense of moral obligation to act consistently with one's values. This means that the process by which Conformism values moderate the influence of Universalism values on cooperative behavior could be expected to involve moral obligation as a mediator variable.

There are two possible scenarios that may produce mediated moderation. In the first scenario, Conformism values could moderate the influence of Universalism values on moral obligation. As argued above, one could expect that in the group of people high in Conformism values, Universalism values would not give rise to a sense of moral obligation, and, therefore, would not influence cooperative behavior. But it does not need to be so. The second possible scenario is that Conformism moderates the influence of moral obligation on cooperative behavior. In either scenario, the connection between Universalism values and cooperative behavior would be severed for those high in Conformism values. In Study 3 we seek to establish whether either or both of these two processes are responsible for the overall moderation of the influence of Universalism values on cooperative behavior.

Study 3 was conducted on the internet. Recent research in the field of personality psychology suggests that questionnaires administered on the web can be trusted (e.g., BUCHANAN, JOHNSON, and GOLDBERG, 2005; GOSLING, VAZIRE, SRIVASTAVA, and JOHN, 2004). However, the advantage of reaching a larger subject pool with a broader range of demographic background variables comes at the cost of losing some control of the subject pool and the experimental situation itself. Concerning the subject pool, although some technical apparatus and skills are required for participation, these are nowadays widely spread in households. Thus, the selection bias in internet studies is not likely to be more serious than in laboratory experiments, as the latter are usually conducted with university sophomores (see also LÖNNQVIST, PAUNONEN, VERKASALO, TUULIO-HENRIKSSON, and LÖNNQVIST (2007) for the effects of personality traits on research participation). Another issue is that on the internet, researchers cannot control the environment of the participants and its potential impact on the decision process. This means there is likely to be more random variance in studies conducted on the internet as compared to laboratory studies, such as Study 2. But this also offers

the opportunity to test the critical assumption of consistency among laboratory and field data. Weighing pros and cons, conducting the study on the internet served as an attempt to gather evidence on cooperative behavior on a much broader and more representative scale than would otherwise have been possible.

4.1 Method

Experimental Design and Procedure

Study 3 was conducted using a questionnaire posted on a University of Helsinki website. The questionnaire was open from June 14th to August 1st 2007. The questionnaire consisted of the Portraits Values Questionnaire, a prisoner's dilemma game, and some other measures not related to the present study. Half of the participants played the PDG game and the other half the PDT game (see Study 2 methods). The instructions, currency, and exchange rate were exactly the same as in Study 2.

There are four important differences between the games played in Studies 2 and 3. As mentioned above, each participant played only one version of the prisoner's dilemma game. Second, the order of the value questionnaire and the game was reversed; participants completed the values questionnaire before they played the prisoner's dilemma game. This was to ensure that participants did not infer their values from their behavior, as could arguably have been the case in Study 2. Third, participants were asked, not only how many Taler they transferred (0 - 10 Taler), but also how many Taler they thought that they should have transferred (0 - 10 Taler). As in Study 2, the former is our measure of cooperative behavior, whereas the latter is our measure of experienced sense of moral obligation. Fourth, participants were told right after the questions concerning the prisoner's dilemma game that only 40 people, picked randomly, would actually be paired with another participant and receive a payment. The method of paying only some of the participants is one of the common research methods in experimental economics (e.g., DREHMANN, OECHSSLER, and ROIDER, 2005; GÜTH, SCHMIDT, and SUTTER, 2003), and has been shown not to distort the results (GÜTH, SCHMIDT, and SUTTER, 2007).

Participants

Participants were recruited through Finnish internet discussion sites and university mailing lists. As an incentive to participate, participants were offered their values profiles. The prisoner's dilemma game or the opportunity to earn money was not mentioned in the invitation. 847 people completed at least parts of the questionnaire. However, 113 people failed to complete the values questionnaire (more than two missing values) or the prisoner's dilemma

game, and were thus excluded from further analyses. This left us with a final sample size of 734 people (428 female (58%), mean age 29.9 years ($SD = 10.8$)).

Measures

As in Studies 1 and 2, values are measured with the Portraits Values Questionnaire (SCHWARTZ et al., 2001). Proportional sum variables are computed to control for differences in scale use (VERKASALO et al., 1996).

4.2 Results and Discussion

The descriptives and internal consistency reliabilities (Cronbach's α) of the ten value types are presented in Table 2. These are generally quite similar to those obtained in Study 2, suggesting that the implementation of the study on the internet does not affect measurement adversely.

First, we explore whether the type of game (treatment) affected transfer decisions. A one-way ANOVA ($F(1, 732) = 4.77, p < .05$) reveals that participants cooperated somewhat more in the PDG game ($M = 6.34, SD = 3.24$) than in the PDT game ($M = 5.82, SD = 3.20$). Consistent with this, participants also thought that they should have transferred more when playing the PDG game ($M = 6.55, SD = 3.40$) than when playing the PDT game ($M = 6.09, SD = 3.33$; one-way ANOVA ($F(1, 732) = 3.39, p = .07$)). However, as in Study 2, the type of game does not interact with Conformism or Universalism values (no significant two-way or three-way interactions in the prediction of behavior or moral obligation, all $F < .46$), suggesting that the data from the two games can be collapsed for the purposes of the present research. We have conducted the analyses below both with collapsed and split samples, and as the results are virtually identical, we present only the results for the collapsed sample. The transferred amount is our measure of cooperative behavior and the amount the participant thought he or she should have transferred is our measure of moral obligation. Moral obligation is highly correlated with cooperative behavior ($r = .67, p < .001$).

The correlations between the ten value types and cooperative behavior are shown in the fifth column of Table 2. As expected, Universalism values positively correlate with the transferred amount ($r = .16, p < .001$). Benevolence values, the other self-transcendence values, also correlate with transferred amount ($r = .15, p < .001$). Power and Achievement values, on the opposite side of the circular structure of values, correlate negatively with cooperative behavior ($r = -.10, p < .05$ and $r = -.07, p < .05$). Hedonism values are also negatively correlated with cooperative behavior ($r = -.09, p < .05$). Although only the positive correlation between Universalism and cooperative behavior has been predicted, the other correlations are

consistent with the structure of values that Schwartz' theory proposes, and could have been predicted based on this structure. That is, values close to Universalism, such as Benevolence, would be expected to correlate positively with cooperative behavior, whereas negative correlations would be expected for values opposite to Universalism, such as Power and Achievement.

Value Type	<i>M</i>	<i>SD</i>	α	r_{Coop}	r_{Mor}
Universalism	1.22	0.20	.81	.16**	.11**
Benevolence	1.15	0.18	.68	.15**	.07*
Conformism	0.84	0.24	.72	.01	-.01
Tradition	0.76	0.21	.50	-.05	-.04
Security	0.98	0.20	.68	-.06	-.06
Power	0.72	0.23	.63	-.10*	-.08*
Achievement	0.88	0.24	.86	-.07*	-.02
Hedonism	1.01	0.25	.91	-.09*	-.04
Stimulation	0.91	0.25	.83	-.01	-.02
Self-Direction	1.23	0.19	.61	.02	-.01

Table 2: Descriptive statistics for the 10 value scales and their correlations with cooperative behavior (r_{Coop}) and moral obligation (r_{Mor}) in two multiple choice Prisoner's Dilemma games ($N = 734$; Study 3), * $p < .05$, ** $p < .01$.

The correlations between values and moral obligation are shown in the sixth column of Table 2. These correlations are somewhat weaker than the correlations between values and behavior, but follow the same pattern.

Mediated Moderation

In our analysis of mediated moderation, we follow the procedures recommended by MULLER, JUDD, and YZERBYT (2005). Table 3 shows the results of three regression analyses predicting (a) cooperative behavior with Universalism, Conformism, and their interaction, (b) moral obligation with Universalism, Conformism, and their interaction, and (c) cooperative behavior with Universalism, Conformism, their interaction, moral obligation, and the interaction between Conformism and moral obligation. Below we discuss these three regression analyses.

Predictors	Dependent Variable					
	Cooperation		Moral Obligation		Cooperation	
	β	<i>t</i>	β	<i>t</i>	β	<i>t</i>
Universalism (U)	.16	4.22**	.12	3.20**	.08	2.67**
Conformism (C)	.01	0.34	.02	0.68	-.01	-0.23
U \times C	-.12	-3.22**	-.09	-2.52*	-.05	-1.85
Moral Obligation (MO)					.66	23.78**
MO \times C					-.02	-1.02

Table 3: Results of regression analyses showing mediated moderation, * $p < .05$, ** $p < .01$.

First, hierarchical multiple regression analyses were performed to examine the moderating effects of Conformism values on Universalism values in the prediction of cooperative behavior (second and third column of Table 3). The regression analyses yielded the predicted significant interaction between Conformism and Universalism ($\beta = -.12$, when the interaction term was added to the regression analysis, adjusted R^2 increased from .01 to .02, the F for the model changed from .8.69 to 8.99, F for the change in $R^2 = 9.38, p < .01$). We estimated regression lines at $\pm 1 SD$ for predictors of the equation (AIKEN and WEST, 1991). Figure 3 provides simple regression lines of cooperative behavior as a function of Universalism values at high and low ends of Conformism values. Tests of simple slopes revealed that the association between Universalism and cooperative behavior was moderate when Conformism values were low, $B = 4.01, SE = 0.77, \beta = .26, t(3,730) = 5.87, p < .001$, but vanished when Conformism values were high, $B = 0.92, SE = 0.82, \beta = .05, t(3,730) = 1.00, p = .27$. These simple slopes suggest that Universalism values guide cooperative behavior only when Conformism values are low.

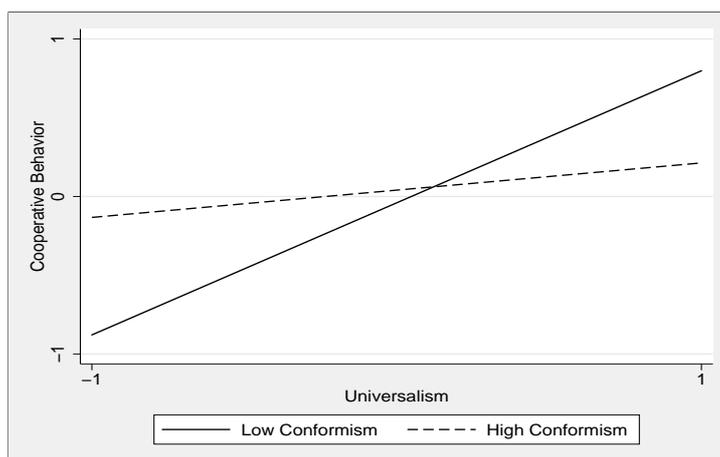


Figure 3: Cooperative behavior in a Prisoner's Dilemma game regressed on low and high Universalism scores for low and high Conformism groups (Study 3).
 Low score = 1 SD below the mean; high score = 1 SD above the mean.

Our second regression analysis regressed moral obligation upon standardized Conformism, Universalism, and their interaction (fourth and fifth columns of Table 3). This regression also yielded the predicted significant interaction between Conformism and Universalism ($\beta = -.09, p < .05$). We estimated regression lines at $\pm 1 SD$ for predictors of the equation. Tests of simple slopes revealed that the association between Universalism and moral obligation was moderate when Conformism values were low, $B = 2.91, SE = 0.82, \beta = .20, t(3,736) = 4.52, p < .001$, but vanished when Conformism values were high, $B = 0.89, SE = 0.86, \beta = .04, t(3,736) = 0.72, p = .45$, suggesting that Universalism values influenced moral obligation only

when Conformism values were low. Conformism values did not moderate the effect of any other values on cooperative behavior or moral obligation.

The above analyses have established that Conformism values moderate the effect of Universalism values on cooperative behavior and on moral obligation. Next, we look at whether the mediating process accounts for the moderation. If it does, then the moderation of the residual direct effect of Universalism values should be reduced compared to the moderation of the overall effect of Universalism. There are two different forms of mediated moderation (see MULLER et al., 2005). The effect of the initial variable (Universalism) on the mediator (moral obligation) may differ as a function of the moderator (Conformism), or the mediator (moral obligation) may interact with the moderator (Conformism) to cause the outcome.

The sixth and seventh columns of Table 3 show the results of our third regression analysis, which regresses cooperative behavior on Universalism, Conformism, their interaction, moral obligation, and the interaction between Conformism and moral obligation. Comparison between the first and third regression analysis presented in Table 3 reveals that the predictive power of the Conformism \times Universalism interaction term decreases from $\beta = -.12(p < .01)$ to $\beta = -.05$ (n.s.) when moral obligation is allowed into the path between the interaction term and cooperative behavior. This shows that once the mediator is controlled, the residual direct effect of Universalism no longer depends on Conformism values. The more conservative Sobel test for mediation effects (SOBEL, 1982) confirms that mediation occurs (Sobel statistic = 2.49, $p < .05$). The third regression analysis in Table 3 also shows that Conformism values do not interact with the moral obligation in predicting cooperative behavior, $\beta = -.05$ (n.s.). This means that Conformism does not moderate the influence of moral obligation on cooperative behavior. Conformism values do also not moderate the influence of any other values on cooperative behavior. Our analyses show that the link between Universalism values and cooperative behavior is severed in the group of people high in Conformism values because their Universalism values do not give rise to a feeling of moral obligation to act consistently with those values.

5 General Discussion

Several theorists have suggested that the more normative or popular a behavior, the weaker its connection to internal dispositions such as personality and values (e.g., JONES and DAVIS, 1965; BARDI and SCHWARTZ, 2003). The reason for this is that popular behaviors, being more socially desirable, are more likely to be guided by social forces than by personal norms. LÖNNQVIST et al. (2006) present an important qualification to this phenomenon: it is true

primarily of people high in Conformism values. The present work extends research on the moderating effect of Conformism values by examining the processes through which Conformism values moderate the influence of other personal values. The results of Study 1 show that Conformism values moderate the influence of group norms on behavior. While people high in Conformism values move toward the group norm, the group norm has no influence on the behavioral intentions of those low in Conformism values. The results of Study 2 show that Universalism values strongly predict cooperative behavior in a one-shot prisoner's dilemma game with multiple choice. However, this is only true in the group of people low in Conformism values; among those high in Conformism values, the relation between Universalism values and cooperative behavior is strongly attenuated. Study 3 expands upon the results of Study 2 by showing that the moderating effect of Conformism values is mediated by experienced sense of moral obligation. That is, Universalism values give rise to an experienced sense of moral obligation and moral obligation influences cooperative behavior. However, this only holds if Conformism values are low.

In Study 1, group norm influences only public, not private behavioral intentions. This may give some clues to the psychological underpinning of the phenomenon. Descriptive norms can exert informational and normative influence. Informational influence is assumed to cause genuine attitude change or internalization of the norm and, therefore, changes both in private and public behaviors. Normative influence, in contrast, does not imply genuine attitude change, but rather a strategic effort to be accepted and avoid social punishment. Consistent with the reasoning of LÖNNQVIST et al. (2006), our results thus imply that Conformism values are connected to fear of social sanctions. This fear may motivate those high in Conformism values to conform to group norms in their display of behaviors. Our results are also in line with previous results demonstrating that those who have a strong personal basis for their attitudes generally do not conform to social norms in their display of behaviors (e.g., HORNSEY et al., 2003).

In the introduction, we noted that influence of personal values on behavior has been debated. Similarly, previous studies on the influence of social norms on behavior have produced somewhat mixed results. Some have seen norms as central to the proper understanding of human behavior (e.g., FISHBEIN and AJZEN, 1975; TRIANDIS, 1977) whereas others have seen little value in the concept (e.g., DARLEY and LATANÉ, 1970; KREBS and MILLER, 1985). This controversy has, in part, been clarified by research on the influence of social norm on behavior that has focused on properties of the norm (e.g., CIALDINI, KALLGREN, and RENO, 1991; CIALDINI et al., 1990). Our results further contribute to this discussion by focusing on properties of the individual. That is, social norms do influence individual behavior, but

primarily in the group of people who consider Conformism values to be relatively important.

In Study 2, participants played a prisoner's dilemma game with multiple choice. This game was probably a novel experience for most or all of our participants. Therefore, one would not expect participants to have knowledge of any social norms on how to behave in such a game. Both the novelty of the situation and the apparent lack of social norms can be expected to strengthen the influence of internal dispositions, such as personal values, on the decisions the participants made. Indeed, Universalism values predict the cooperative behavior of all our participants, but much more strongly for those people who are low in Conformism values. The data of Study 2 is the first to show that Conformism values moderate the influence of other personal values on individual behavior. This result is even more impressive considering that research on the predictability of behavior (e.g., DIENER and LARSEN, 1984; EPSTEIN, 1979; MOSKOWITZ, 1982) suggests that, although the mean of the distribution of behavior is often predictable, single behaviors are among the least predictable variables in psychology. Furthermore, the results of Studies 2 and 3 suggest that Conformism values moderate the influence of Universalism values on different types of cooperative behavior. That is, whether cooperative behavior means that one refrains from anti-social behavior (transferring money from the other person to oneself) or that one engages in pro-social behavior (transferring money from oneself to the other person), Conformism values moderate the influence of Universalism values on that behavior.

Our focus on moral obligation in Study 3 provides us with some insight concerning the processes by which Universalism values influence behavior. As expected, Universalism values give rise to a sense of moral obligation to act consistently with those values. This sense of moral obligation mediates the effect of Universalism values on cooperative behavior. More to the point of the present discussion, Universalism values are associated with a sense of moral obligation only in the group of people low in Conformism values. The results show that the link between Universalism values and cooperative behavior is severed among those high in Conformism values because the Universalism values of these people do not give rise to a sense of moral obligation.

One difference between Universalism values and other values (e.g., Stimulation, Security) is that the influence of other values on behavior might not be as strongly mediated by experienced sense of moral obligation (see LÖNNQVIST et al., 2006). This could, in part, explain why Conformism values may primarily moderate the influence of Universalism values, but perhaps not other values, on behavior. A recent review suggests that the level of moral reasoning, as conceptualized through KOHLBERG'S (1984) stages, is inversely related to Con-

formism values (HELKAMA, 2004). People at the lower or conventional stages of morality think as members of the conventional society with its values and norms. This means that people high in Conformism values may feel strongly morally obliged by the values and norms that society provides. At the higher or postconventional stages of morality, people make more independent efforts to think out what any society ought to value. This means that those low in Conformism values may feel more strongly morally obliged by their personal values. At this point we want to emphasize that the results of our research in no way indicate that people who value Conformism behave less cooperatively than people who do not value Conformism. Indeed, Conformism in itself is not correlated with cooperative behavior or moral obligation in either of the prisoner's dilemma games.

TRIANDIS (1995) has argued that social roles and norms are more important in collectivistic than individualistic cultures and that the influence of internal dispositions on behavior is, therefore, weaker in collectivistic cultures. It is possible that the relations between other personal values and behavior are weaker in cultures where Conformism values are generally considered important. Thus, the influence of social norms on behavior may be more pronounced in cultures where Conformism values are considered more important. Although any generalizations from individual-level differences to culture-level differences must remain speculative, cross-cultural differences may be worthy of exploration in future studies.

Finally, some general limitations of the present research ought to be acknowledged. First, the moderating effects of Conformism values were generally rather weak. In Studies 1 and 2, conducted with small samples, the interactive effects were only marginally statistically significant, and the results must therefore be interpreted with caution. Although Study 3, conducted with a much larger sample, confirmed the results of Study 2, it should be noted that the interaction effect in Study 3 accounted for a rather small increase in explained variance. Thus, although the effects we found can be considered theoretically important, their magnitude should not be exaggerated. Another important limitation was that the measure of moral obligation used in Study 3 was probably not the best possible index of that construct. In particular, a more sophisticated, multi-item measure would have been preferable and should be used in future research. Furthermore, it might have been instructive to look for possible order effects in Study 3. In the present research design, however, this was not possible as all participants responded first to the behavioral measure, and then to the moral obligation measure.

In conclusions, our results provide us with new information on how and when personal values and social norms predict behavior. Both values and norms are concepts whose role in the explanation of behavior has frequently been contested (e.g., DARLEY and LATANÉ, 1970;

KREBS and MILLER, 1985; KRISTIANSEN and HOTTE, 1996; MCCLELLAND, 1985). Previous research on the influence of personal values and social norms on behavior has clarified this issue by showing that norms, whether internal or external, will only guide behavior when they are focal at the time of the behavior (KALLGREN, RENO, and CIALDINI, 2000; VERPLANKEN and HOLLAND, 2002). Our results and the results of LÖNNQVIST et al. (2006) suggest that there exist individual differences in the extent to which individuals naturally focus on internal or external norms, and that the individuals' Conformism values are one determinant of these individual differences. The present research establishes (a) that social norms can influence people high in Conformism values to a degree that obscures the true influence of attitudes or values on behavior, (b) that Universalism values predict individual cooperative behavior only for people who are low in Conformism values, and (c) that the reason for this is that Universalism values give rise to a sense of moral obligation only in those people.

References

- ABRAMS, D., M. WETHERELL, S. COCHRANE, M.A. HOGG, and J.C. TURNER (1990): "Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization", *British Journal of Social Psychology*, 29(4), 97-119.
- AIKEN, L.S. and S.G. WEST (1991): *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- AJZEN, I. (1991): "The theory of planned behavior", *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.
- ASCH, S.E. (1952): "Effects of group pressure on the modification and distortion of judgments", in G.E. SWANSON, T.M. NEWCOMB and E.L. HARTLEY (Eds.), *Readings in social psychology*, New York: Holt, Reinhart, & Winston, pp. 393-401.
- BARDI, A.B. and S.H. SCHWARTZ (2003): "Value and behavior: Strength and structure of relations", *Personality and Social Psychology Bulletin*, 29(10), 1207-1220.
- BECK, L. and I. AJZEN (1991): "Predicting dishonest actions using the theory of planned behavior", *Journal of Research in Personality*, 25(3), 285-301.
- BERNARD, M.M., G.R. MAIO, and J.M. OLSON (2003): "The vulnerability of values to attack: Inoculation of values and value-relevant attitudes", *Personality and Social Psychology Bulletin*, 29(1), 63-75.
- BUCHANAN, T., J.A. JOHNSON, and L.R. GOLDBERG (2005): "Implementing a five-factor personality inventory for use on the Internet", *European Journal of Psychological Assessment*, 21(2), 116-128.

- CIALDINI, R.B., C.A. KALLGREN, and R.R. RENO (1991): "A focus theory of normative conduct: a theoretical refinement and reevaluation of the role of norms in human behavior", *Advances in Experimental Social Psychology*, 24, 201-234.
- CIALDINI, R.B., R.R. RENO, and C.A. KALLGREN (1990): "A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places", *Journal of Personality and Social Psychology*, 58, 1015-1026.
- DARLEY, J.M. and C.D. BATSON (1973): "From Jerusalem to Jericho: A study of situational and dispositional variables in helping behavior", *Journal of Personality and Social Psychology*, 27(1), 100-108.
- DARLEY, J.M. and B. LATANÉ (1970): "Norms and normative behavior: Field studies of social interdependence", in J. MACAULAY and L. BERKOWITZ (Eds.), *Altruism and helping behavior*, New York: Academic Press, pp. 83-102.
- DIENER, E. and R.J. LARSEN (1984): "Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses", *Journal of Personality and Social Psychology*, 47(4), 871-883.
- DREHMANN, M., J. OECHSSLER, and A. ROIDER (2005): "Herding and contrarian behavior in financial markets: An internet experiment", *American Economic Review*, 95(5), 1403-1426.
- EPSTEIN, S. (1979): "The stability of behavior: I. On predicting most of the people much of the time", *Journal of Personality and Social Psychology*, 37, 1097-1126.
- FISHBEIN, M. and I. AJZEN (1975): *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- GOERG, S. and G. WALKOWITZ (2008): "Presentation Effects in Cross-Cultural Experiments - An Experimental Framework for Comparisons", Bonn Graduate School of Economics, University of Bonn: *Discussion Paper No. 5/2008*.
- GOSLING, S.D., S. VAZIRE, S. SRIVASTAVA, and O.P. JOHN (2004): "Should we trust web-based studies? A comparative analysis of six preconceptions about Internet questionnaires", *American Psychologist*, 59(2), 93-104.
- GUIMOND, S., N.R. BRANSCOMBE, S. BRUNOT, A.P. BUUNK, A. CHATARD, M. DÉSERTE, D.M. GARCIA, S. HAQUE, D. MARTINOT, and V. YZERBYT (2007): "Culture, gender, and the self: Variations and impact of social comparison processes", *Journal of Personality and Social Psychology*, 92(6), 1118-1134.
- GÜTH, W., C. SCHMIDT, and M. SUTTER (2003): "Fairness in the mail and opportunism in the Internet - A newspaper experiment on ultimatum bargaining", *German Economic Review*, 4(2), 243-265.

- GÜTH, W., C. SCHMIDT, and M. SUTTER (2007): "Bargaining outside the lab - A newspaper experiment of a three-person ultimatum game", *The Economic Journal*, 117(518), 442-448.
- HELKAMA, K. (2004): "Values, role-taking, and empathy in moral development", *New Review of Social Psychology*, 3(1-2), 103-111.
- HORNSEY, M.J., L. MAJKUT, D.J. TERRY, and B.M. MCKIMMIE (2003): "On being loud and proud: Non-conformity and counter-conformity to group norms", *British Journal of Social Psychology*, 42(3), 319-335.
- INSKO, C., S. DRENAN, M.R. SOLOMON, R. SMITH, and T.J. WADE (1983): "Conformity as a function of the consistency of positive self-evaluation with being liked and being right", *Journal of Experimental Social Psychology*, 19(4), 341-358.
- JACKSON, D.N. and S.V. PAUNONEN (1985): "Construct Validity and the Predictability of Behaviour", *Journal of Personality and Social Psychology*, 49, 554-570.
- JONES, E.E. and K.E. DAVIS (1965): "From acts to dispositions: the attribution process in social psychology", in L. BERKOWITZ (Ed.), *Advances in experimental social psychology*, New York: Academic Press, Vol. 2, pp. 219-266.
- KALLGREN, C.A., R.R. RENO, and R.B. CIALDINI (2000): "A focus theory of normative conduct: When norms do and do not affect behavior", *Personality and Social Psychology Bulletin*, 26, 1002-1012.
- KOHLBERG, L. (1984): *The psychology of moral development: The nature and validity of moral stages*. San Francisco: Harper & Row.
- KREBS, D.L. and D.T. MILLER (1985): "Altruism and aggression", in G. LINDZEY and E. ARONSON (Eds.), *The handbook of social psychology*, Third Edition, Vol. 2, New York: Random House, pp. 1-71.
- KRISTIANSEN, C.M. and A.M. HOTTE (1996): "Morality and the self: Implications for the when and how of value-attitude-behavior relations", in C. SELIGMAN, J.M. OLSON, and M.P. ZANNA (Eds.), *The Psychology of Values*, Mahwah, NJ: Erlbaum, pp. 77-105.
- LÖNNQVIST, J-E., S. LEIKAS, S.V. PAUNONEN, V. NISSINEN, and M. VERKASALO (2006): "Conformism moderates the relations between values, anticipated regret, and behavior", *Personality and Social Psychology Bulletin*, 32(11), 1469-1481.
- LÖNNQVIST, J-E., S.V. PAUNONEN, M. VERKASALO, S. LEIKAS, A. TUULIO-HENRIKSSON, and J. LÖNNQVIST (2007): "Personality characteristics of research volunteers", *European Journal of Personality*, 21, 1017-1030.
- MAIO, G.R. and J.M. OLSON (1995): "Relations between values, attitudes, and behavioral intentions: the moderating role of attitude function", *Journal of Experimental Social Psychology*, 31, 101-115.

- Psychology*, 31(3), 266-285.
- MAIO, G.R., J.M. OLSON, L. ALLEN, and M.M. BERNARD (2001): "Addressing discrepancies between values and behavior: The motivating effects of reasons", *Journal of Experimental Social Psychology*, 17, 104-117.
- MAIO, G.R., J.M. OLSON, M.M. BERNARD, and M.A. LUKE (2003): "Ideologies, values, attitudes, and behavior", in J. DELAMATER (Ed.), *Handbook of social psychology*, New York: Kluwer Academic, pp. 283-308.
- MCCLELLAND, D.C. (1985): *Human Motivation*. Glenview, IL: Scott, Foresman, and Company.
- MCCLELLAND, D.C., R. KOESTNER, and J. WEINBERGER (1989): "How do self-attributed and implicit motives differ?", *Psychological Review*, 96, 690-702.
- MISCHEL, W. and P.K. PEAKE (1982): "Beyond déjà vu in the search for cross-situational consistency", *Psychological Review*, 89, 730-755.
- MOSKOWITZ, D.S. (1982): "Coherence and cross-situational generality in personality: A new analysis of old problems", *Journal of Personality and Social Psychology*, 43, 754-768.
- MULLER, D., C.M. JUDD, and V.Y. YZERBYT (2005): "When moderation is mediated and mediation is moderated", *Journal of Personality and Social Psychology*, 89, 852-863.
- MURRAY, H.A. (1938): *Explorations in personality*. New York: Oxford University Press.
- PAULHUS, D. L. and O.P. JOHN (1998): "Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives", *Journal of Personality*, Special Issue: Defence Mechanisms in Contemporary Personality Research, 66, 1025-1060.
- ROKEACH, M. (1973): *The nature of human values*. New York: Free Press.
- ROSS, L., and R.E. NISBETT (1991): *The person and the situation: Perspectives of social psychology*. New York: McGraw-Hill.
- SCHWARTZ, S. (1977): "Normative influences on altruism", in L. BERKOWITZ (Ed.), *Advances in experimental social psychology*, San Diego, CA: Academic Press, Vol. 10, pp. 221-279.
- SCHWARTZ, S.H. (1992): "Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries", in M.P. ZANNA (Ed.), *Advances in experimental social psychology*, New York: Academic Press, Vol. 25, pp. 1-65.
- SCHWARTZ, S.H. (1996): "Value priorities and behavior: Applying a theory of integrated value systems", in C. SELIGMAN, J.M. OLSON, and M.P. ZANNA (Eds.), *The psychology of values: The Ontario Symposium*, Hillsdale, NJ: Erlbaum, Vol. 8, pp. 1-24.

- SCHWARTZ, S.H., G. MELECH, A. LEHMANN, S. BURGESS, M. HARRIS, and V. OWENS (2001): "Extending the cross-cultural validity of the theory of basic human values with a different method of measurement", *Journal of Cross-Cultural Psychology*, 32, 519-542.
- SCHWARTZ, S.H. (2005a): "Basic human values: Their content and structure across countries", in A. TAMAYO and J.B. PORTO (Eds.), *Valores e Comportamento nas Organizações* [Values and Behavior in Organizations], Petrópolis, Brazil: Vozes, pp. 21-55.
- SCHWARTZ, S.H. (2005b): "Robustness and fruitfulness of a theory of universals in individual human values", in A. TAMAYO and J.B. PORTO (Eds.), *Valores e Comportamento nas Organizações* [Values and Behavior in Organizations], Petrópolis, Brazil: Vozes, pp. 56-95.
- SCHWARTZ, S.H. and L. SAGIV (1995): "Identifying culture specifics in the content and structure of values", *Journal of Cross-Cultural Psychology*, 26, 92-116.
- SCHWARTZ, S.H. and R.C. TESSLER (1972): "A test of a model for reducing measured attitude-behavior discrepancies", *Journal of Personality and Social Psychology*, 24, 225-236.
- SHERIF, M. and C.W. SHERIF (1964): Reference groups: New York: Harper & Row.
- SOBEL, M. E. (1982): "Asymptotic confidence intervals for indirect effects in structural equation models", in S. LEINHARDT (Ed.), *Sociological Methodology*, Washington DC: American Sociological Association, pp. 290-312.
- TRIANDIS, H.C. (1977): Interpersonal behavior. Monterey, CA: Brooks/Cole.
- TRIANDIS, H.C. (1995): Individualism and collectivism. Boulder, CO: Westview Press.
- VERKASALO, M., P. TUOMIVAARA, and M. LINDEMAN (1996): "15-year-old pupils' and their teachers' values, and their beliefs about the values of an ideal pupil", *Educational Psychology*, 16, 35-47.
- VERPLANKEN, B. and R.W. HOLLAND (2002): "Motivated decision-making: Effects of activation and self-centrality of values on choices and behavior", *Journal of Personality and Social Psychology*, 82, 434-447.

Appendix: Schwartz' value types (SCHWARTZ, 1992)

Benevolence: Those who tend towards benevolence are very giving, seeking to help others and provide general welfare. They are the 'earth mothers' who nurture all.

Tradition: The traditionalist respects that which has gone before, doing things simply because they are customary. They are conservatives in the original sense, seeking to preserve the world order as is. Any change makes them uncomfortable.

Conformity: The person who values conformity seeks obedience to clear rules and structures. They gain a sense of control through doing what they are told and conforming to agreed laws and statutes.

Security: Those who seek security seek health and safety to a greater degree than other people. Though they may worry about the potential of military force, they welcome the comfort that their existence brings.

Power: This takes value from social status and prestige. The ability to control others is important and power will be actively sought through dominance of others and control over resources.

Achievement: Value here comes from setting goals and then achieving them. The more challenge, the greater the sense of achievement. When others have achieved the same thing, status is reduced and greater goals are sought.

Hedonism: Hedonists simply enjoy themselves. They seek pleasure above all things and may, according to the view of others, sink into debauchery.

Stimulation: The need for stimulation is close to hedonism, though the goal is slightly different. Pleasure here comes more specifically from excitement and thrills and a person with this driver is more likely to be found doing extreme sports than propping up a bar.

Self-direction: Those who seek self-direction enjoy being independent and outside the control of others. They prefer freedom and may have a particular creative or artistic bent, which they seek to indulge whenever possible.

Universalism: The universalist seeks social justice and tolerance for all. They promote peace and equality and find war anathema except perhaps in pursuit of lasting peace.

Chapter III

**Actions and Beliefs in a Trilateral Trust Game Involving
Germans, Israelis, and Palestinians***

*“To be trusted is a greater compliment
than to be loved.”*

GEORGE MACDONALD

1 Introduction

Recent papers in experimental economics discuss cross-cultural¹ experimental results that compare findings on behavioral regularities across societies and test their global validity (e.g., ROTH, PRASNIKAR, OKUNO-FUJIWARA, and ZAMIR, 1991; HENRICH, BOYD, BOWLES, CAMERER, FEHR, GINTIS, and MCELREATH, 2001; BUCHAN and CROSON, 2004; BUCHAN, CROSON, and JOHNSON, 2004; HOLM and DANIELSON, 2005; LAZZARINI, MADALAZZO, ARTES and SIQUEIRA, 2005; CHUAH, HOFFMANN, JONES, and WILLIAMS, 2007; HERRMANN, THÖNI, and GÄCHTER, 2008). These studies revealed behavioral similarities and differences across countries and cultures (e.g., in social preferences or cooperative behavior).

Participants from these experiments are typically drawn from the same country or culture. Direct inter-cultural exchange, however, might be governed by quite different behavioral determinants than mere intra-cultural interactions. Thus, inter-cultural experimental studies focusing on interactive strategic decision making across national borders are called for to explore this topic. The present chapter contributes to this research agenda by a controlled trilateral inter-cultural trust experiment involving Germans, Israelis, and Palestinians.

*This chapter is based on the paper: “Actions and Beliefs in a Trilateral Trust Game Involving Germans, Israelis, and Palestinians” by HEIKE HENNIG-SCHMIDT, REINHARD SELTEN, GARI WALKOWITZ, and EYAL WINTER; submitted manuscript, under review.

¹We distinguish between intra-cultural, cross-cultural, and inter-cultural experiments (WALKOWITZ, OBERHAMMER, and HENNIG-SCHMIDT, 2004). In *intra-cultural* experiments, participants from one country/culture interact with each other. *Cross-cultural* experiments compare behavior in two or more countries. *Inter-cultural* experiments investigate behavior of subjects in different countries interacting directly with each other.

Surprisingly, inter-cultural experimental investigations are rare² although globalized labor and commodity markets increasingly lead to everyday interactions of multi-national groups of subjects holding diverse cultural backgrounds and affiliations. Not knowing how culture-specific attitudes influence exchange among foreigners can lead to mutual misunderstanding, frictions, and conflicts not primarily intended. Moreover, to be aware of similarities might strengthen cross-border relationships and lead to more efficient social and economic exchange.

Ingroup consciousness, cross-cultural comparisons of standards and norms, and discrimination are the subjects of the ethnocentrism literature (see SUMNER, 1906; ADAMS, 1951; SCHOPMEYER and FISHER, 1993; KALIN and BERRY, 1994). The ENCYCLOPAEDIA BRITANNICA (2008) defines ethnocentrism as the tendency to interpret or evaluate other cultures in terms of one's own (see also KALIN and BERRY, 1994). Furthermore, ethnocentrism contains that individuals typically attribute the standards of behavior that exist in their own environment also to foreigners. This might be even enforced by the fact that individuals are subconsciously motivated to persist in their beliefs. They attend to evidence and arguments in a selective fashion, undermining them (KAHAN, 2008). A further characteristic of ethnocentrism is the affinity to consider one's own cultural norms and values as universally valid and superior to all the others (THE COLUMBIA ENCYCLOPEDIA, 2007). This implies that members of one cultural group have an exaggerated preference for their own group and an accompanying dislike of other cultural groups which might result in discriminative attitudes toward representatives of these particular other groups. Consequently, flaunt discrimination can be one source of inter-cultural conflict. Taken together, these pieces of evidence might lead to a false perception of the intentions underlying foreigners' behavior. In case of well-pronounced cultural differences a foreigner's decision might be perceived as discriminative or hostile, that in fact is not but has to be attributed to the different behavioral standards and norms prevailing in the foreign culture. Different standards and norms are particularly likely if cultures vary in their religious heritage (see INGLEHART, BASAÑEZ, and MORENO, 2001; TAN and VOGEL, 2005) as in Germany, Israel, and Palestine³. This discord bears huge potential for conflicts as distrusting behavior can be triggered and mutually enforced.

Why is trust a key element in an inter-cultural environment? Trust - as well as trustworthiness⁴ - focuses on common interests in human interactions. They display constitutive mechanisms in almost every bilateral work relationship and economic transaction and re-

²See for instance BOARINI, LASLIER, and ROBIN, 2006; WILLINGER, KESER, LOHMANN, and USUNIER, 2003; WALKOWITZ et al., 2004; CASTRO, 2007; GOERG, GÜTH, HENNIG-SCHMIDT, WALKOWITZ, and WEILAND, 2008.

³At the moment, a Palestinian state does not exist. Most of our Palestinian subjects are formally citizens of Jordan and Israel. For the sake of simplifying the notation, we, nevertheless, refer to them as Palestinians.

⁴In the following, we use the terms "trustworthiness" and "reciprocity" as synonyms.

duce social complexity (ARROW, 1972; LUHMANN, 1979). For instance, in new, unknown, or unstable foreign markets trustful relationships can create competitive advantages that enable a better implementation of strategies for multi-national companies. Lower transaction and agency costs arise as parties need not invest huge amounts of money to protect themselves against opportunistic moves of the exchange partner (WILLIAMSON, 1981; FUKUYAMA, 1995). Moreover, higher trust is associated with stronger economic performance as well as with higher and more equally distributed incomes (KNACK and KEEFER, 1997)⁵ - on the organizational level as well as within the scope of an entire economy.

To study the concept of ethnocentrism and its impact on trust and trustworthiness in an actual inter-cultural environment we apply a trilateral trust game experiment based on the investment game by BERG, DICKHAUT, and MCCABE (1995) as our workhorse. Subjects from Germany, Israel, and Palestinians from the Westbank act either in the sender or in the responder role. The religious, historical, and political background of the participating subject pools makes them a promising testbed for investigating culturally induced behavioral differences and potential ethnocentric attitudes. In fact, our experiment does not only involve participants with different religious backgrounds - Christianity for Germans, Judaism for Israelis, and Islam for Palestinians - but also representatives from present political conflict parties (Israelis and Palestinians) and with a difficult joint historical past (Germans and Israelis).

Our trilateral trust game is a modified version of a sequential two-player investment game where a sender⁶ can transfer any non-negative part a of his endowment to an anonymous responder. The responder pockets her endowment. While being transferred, a is tripled by the experimenters. The responder finally decides to back-transfer any non-negative amount b to the sender. In our setup, each participant sequentially plays the trust game with one member of his/her own national group as well as with one member of each of the other two subject pools in a randomized order. Accordingly, senders in all three subject pools send transfers to three culturally different responders. Responders also decide how much to back-transfer to senders located in all three subject pools. Consequently, a player's actions always consist of one intra-cultural interaction and two interactions with two different foreign players. In addition to transfers and back-transfers, we elicit senders' and responders' first order beliefs on the matched players' actions.

⁵For the impact of culture and social background on attitudes and behavior see the large-scale survey studies of SCHWARTZ, 1992; PUTNAM, LEONARDI, NANETTI, 1994; FUKUYAMA, 1995; HOFSTEDE, 2001; INGLEHART et al., 2001; and MACKIE, 2001.

⁶In the following, we denote the sender as male and the responder as female.

Under perfect information, the standard (albeit Pareto-inefficient) subgame perfect as well as the Nash equilibrium prediction of this game is that rational senders should transfer nothing, since they realize that a purely self-interested responder has no incentive to send any amount back to them. Transfers and back-transfers are typically regarded as measures of trust and trustworthiness, respectively. A sender's positive transfer in general reflects his expectation about the degree the responder is willing to deviate from the subgame perfect solution and to act reciprocally. Responder's back-transfers typically exhibit her willingness to reward or reciprocate sender's trust⁷. Virtually all intra-cultural studies conducted so far have strongly rejected the standard game theoretic prediction (see e.g., CARDENAS and CARPENTER, 2004a). Senders typically transfer about 50% of their initial endowment; responders choose to back-transfer roughly 40% of the received amount.

Our experimental results match very well the findings of the literature on cultural differences and ethnocentrism in a way that different cultural standards exist and players tend to evaluate other cultures in terms of their own. We find levels of trust and beliefs on trust to differ substantially in the three subject pools with Palestinians sending the highest and Israelis sending the lowest transfer amounts. Moreover, responders' beliefs match actual trust behavior of senders from the own country very well. In addition, senders in general do not discriminate in their beliefs toward fellow-countrymen and foreigners. However, the most important result of this chapter is that subjects tend to attribute the standards of behavior prevailing in their own environment also to participants of the other countries thereby ignoring cultural, behavioral, and attitudinal differences between countries. Yet, they are completely wrong in predicting the behavior of players in the other countries. In our case this might lead to disappointment, specifically of Palestinians, and positive surprise for Israeli subjects. Contrary to what might be also expected from the ethnocentrism literature, we find only minor discrimination in actual trusting behavior.

Results on reciprocity are not as clear-cut. While the above results hold for actual reciprocity behavior and beliefs of Germans and Israelis, most significant differences are found in interactions and beliefs when Palestinians are involved.

The remainder of the chapter is organized as follows: In the next section 2 we introduce our research questions. The subsequent section 3 gives a detailed description of the experimental design, applied methods, and procedure. Section 4 presents our results on (back-)transfer standards and ethnocentrism in actual decisions and beliefs on trust and trustworthiness as

⁷For an investigation of the underlying motives of transfers and back-transfers in the investment game see JACOBSEN and SADRIEH (1996).

well as an analysis of the outcome of our trilateral experiment. In section 5 we discuss the possible consequences of our findings. In the final section 6 we conclude.

2 Research questions

The first research question we want to address in this chapter is whether observed behavior in our experiment matches the theoretical and empirical literature on culture differences (e.g., ROTH et al., 1991; HOFSTEDE, 2001; INGLEHART et al., 2001). In particular, do trust and trustworthiness differ across cultures and do different cultural standards exist with regard to these variables? We conjecture that due to our participants' different cultural, religious, and political backgrounds and because of the differences in economic performance of their respective points of origin⁸, the observed levels of trust and reciprocity differ substantially among the three subject pools. We, therefore, compare displayed behavior across countries. In addition, we are interested whether comparable culture-specific standards also prevail in subjects' beliefs on related actions. Are intra-cultural trust and trustworthiness standards also reflected in the intra-cultural expectations concerning actual behavior? Thus, in a second step we investigate whether subjects have correct beliefs on actual behavior of fellow-country counterparts.

Our second research question is concerned with an important result of the ethnocentrism literature, namely that subjects tend to evaluate other cultures and their representatives' behavior in terms of their own, considering their home cultural standards and norms as universally valid. Applied to our experiment, these findings would predict that subjects form correct beliefs on actual behavior of fellow-country counterparts but - given cultural differences - that they attribute the same belief to other (foreign) countries. Moreover, within a country, beliefs about the behavior in all three assigned subject pools should not differ substantially. If this is the case behavior might be misinterpreted by foreigners as positive or negative discriminative that in fact is not but has to be attributed to culture-specific behavioral standards and associated expectations. We will first investigate whether attributed beliefs vary dependent on culture affiliation of the matched interaction partner. We then compare intra-cultural actions and beliefs assigned to foreigners' actions. Finally, we will show the implications of our findings.

Our third research question deals with the finding that one's own cultural standards and

⁸In 2000 Gross Domestic Product (GDP) per capita in Germany was 27,000 US\$ and in Israel 15,000 US\$. In the Westbank, GDP per capita amounted to 2,000 US\$ in 1996 (Source: United Nations Statistics Division, <http://unstats.un.org>). We are aware that broad economic differences between the Westbank and the other Palestinian areas exist.

norms tend to be considered superior to all others. If such discriminative behavior has prevailed in our subject pools we should observe that fellow-country participants are dealt with in a different way than foreign counterparts⁹. Therefore, we will investigate whether assigned participants are treated differently dependent on their subject pool affiliation.

The last research question relates to the outcome of our trilateral trust-game. We are especially interested in the monetary consequences of trust and reciprocity in intra-cultural and inter-cultural exchanges. In particular, did trust pay in our trilateral trust experiment? And if so, when and where? For this analysis we calculate the return on investment for all matchings and senders.

All our research questions are examined with regard to trust and reciprocity.

3 Experimental Methods and Procedure

3.1 Experimental methods

The present trust game experiment is a trilateral adaptation of the usual sequential protocol for running investment game experiments (BERG et al., 1995). A sender and a responder are endowed with $X = 10$ ECU (Experimental Currency Unit) each. The responder pockets her endowment. The sender can transfer any non-negative part a of his endowment to the anonymous responder. Each amount a transferred by the sender is tripled by the experimenter and given to the responder. The responder then decides to back-transfer any non-negative fraction b of the received amount $3a$ to the sender.

In our setup, each participant sequentially plays the trust game with one member of his/her own national subject pool as well as with one member of each of the other two subject pools in a randomized order. Thus, senders in all three subject pools send transfers to three culturally different responders. Responders decide how much to back-transfer to senders in all three subject pools. In addition to transfers and back-transfers, we asked senders and responders to state first order beliefs on the matched players' actions.

⁹Discriminative behavior might be caused by several reasons. First, people might just have a taste for discrimination and are willing to sacrifice resources to cater to their prejudice (BECKER, 1957). Further, foreigners might be treated differently because their expected average ability or reciprocity is significantly lower than the one of fellow-countrymen. Thus, adapting to their stereotypes or prejudices people maximize their expected income by discriminating foreigners. This kind of discrimination is also termed "rational" or "statistical" discrimination (ARROW, 1971; PHELPS, 1972; AIGNER and CAIN, 1977). FERSHTMAN and GNEEZY (2001) describe ethnic discrimination as the outcome of these two aspects. They add the possibility of people having stereotypes about other cultures that can be either true or not. In the case that these stereotypes are true, we have the special case of statistical discrimination. Even if they are not true, they may influence the interaction between individuals or groups and lead to biased decisions.

A trilateral multi-person multi-stage experiment like ours encounters substantial organizational efforts due to the geographically distant locations of the three involved subject pools. The crucial constraint is the interdependence of sequential choices made by senders and responders. The responder in one country cannot make a direct cross-border back-transfer without knowing the respective foreign-country sender's transfer. We solve this problem by combining several experimental features known from the literature or developed by ourselves. We call our approach *Experimenting over a Long Distance* (ELD). ELD was first applied by WALKOWITZ et al. (2004)¹⁰.

In the experiment reported here, ELD owns the following features:

1. *Strategy method*. The strategy method (SELTEN, 1967) allows to organizationally disconnect the second stage of the game from the first stage. By having the responder state her back-transfer for each possible transfer of the sender, the sequential two-person two-stage game is converted into a two-person normal-form one-stage game for each player. Consequently, by this procedure the game can be played with time delay independently at different locations. For applications of the strategy method see also BRANDTS and CHARNESS (2000), GÜTH, SCHMIDT, and SUTTER (2003), and BELLEMARE and KRÖGER (2007).

2. *Pen-and-paper*. Using pen-and-paper makes the experimental design independent of technical equipment and software compatibility and start-up costs are immensely reduced. This is particularly important in our trilateral context where a common hardware and software standard was not available.

3. *Remote-control organization*. We designed a sophisticated experimental procedure to ensure equivalent experimental conditions for all three subject pools. A central unit, the Chief Experimenter (CE) was located in Bonn. Local Experimenters (LEs) were in charge of organizing and running the sessions in Jerusalem and Bethlehem¹¹. CE instructed all LEs in advance by distributing a detailed procedural script and an extensive instruction manual in English. The translation of the instructions from German into the local languages was done in Bonn by native speakers applying the back-translation method (BRISLIN, 1970; ECO, 2000). CE ex ante coded and prepared the decision sheets and randomly matched players across subject pools. Participants at each location randomly drew a personal identification code constituting a predefined order of matching across subject pools, not noticeable for participants. The code also ensured full anonymity by a double-blind procedure. Subjects

¹⁰The ELD-method is detailly introduced in Appendix D.

¹¹It is advisable in inter-cultural and cross-cultural research to work with local experimenters in each national subject pool in order to avoid self-presentation and face-saving effects (BOND and HWANG, 1986).

then made their choices on decision sheets marked with their code number and displaying their counterparts' pool affiliations. After all sessions in all locations had been finished, CE collected the data, computed the payoffs, and transferred this information to all LEs. Finally, subjects were paid out by LEs two to three days after the last session in any of the subject pools had been finished.

3.2 Experimental procedure

The trilateral experiment was run simultaneously at BonnEconLab (Laboratory for Experimental Economics, University of Bonn, Germany), at RatioLab (Laboratory for the Study of Interactive Decision Making, Hebrew University of Jerusalem, Israel), and in a classroom at Bethlehem University, Palestine. 90 students participated majoring in different disciplines¹² at the above universities. Subjects were recruited by campus advertisements¹³ promising a monetary reward for participation in a decision-making task. We strictly controlled for participants' nationalities in order to exclude foreign players from national subject pools¹⁴ and for age and gender to guarantee similar conditions in each university. In each of the three countries, we ran a sender session and a responder session with 15 participants each¹⁵.

Any German, Israeli, or Palestinian participant played the trust game with one member of his/her own country as well as with one member of the other two countries. All matchings were played one-shot. Thus, each player interacted with one subject of every of the three national subject pools once, making choices in three subsequent decision rounds. The order of decisions was randomized beforehand. In each round, any player was endowed with $X=10$ ECU¹⁶ with each sender making a transfer a and each responder deciding on a back-transfer b . Participants kept their player type throughout the experiment. A player's total payoff was the sum of the payoffs over all three decision rounds plus the show-up fee.

When introducing the experiment, we informed subjects about the countries involved in the study by providing them with the names of the participating universities and the countries they are located in. Due to religious, organizational, and technical constraints, the experiment could not be run at exactly the same time at all locations. We, therefore, conducted sessions

¹²Economics, Mathematics, Law studies, Sociology, Psychology, Philosophy, Business administration, and Medicine.

¹³In Bonn, subjects were recruited by the online recruiting system ORSEE (GREINER, 2004).

¹⁴In Jerusalem, Israeli Arab students did not participate in the experiment.

¹⁵We had to restrict the number of independent observations to 15 due to limited availability of subjects at Bethlehem University. Moreover, we confined our analysis to one university per country as we were not able to find a second Palestinian university to be included into our study.

¹⁶The exchange rate was 1 ECU=0.50 US\$ in each population. As the cross-national decisions were highly interdependent we did not want to adjust the exchange rate to local purchasing power. We paid a locally adapted show up fee instead (5 US\$ for Germans and Israelis, 3 US\$ for Palestinians).

on three consecutive days. At each location, LEs ran the sender and responder sessions within one day.

The instructions of the trilateral trust game were read aloud and explained in detail to the experimental subjects¹⁷. Questions were answered in private. We took great care to ensure participants' understanding of the game and the underlying incentive structure. After having read the instruction again, for which they had plenty of time, subjects had to answer four control questions testing their understanding of the decision task. We did not proceed in the experimental protocol until all subjects had answered all questions correctly. Thus, we can safely assume that people fully understood the game, its features, and incentives.

After having completed the control questions, senders had to make their transfer choices a and responders had to write down their back-transfer decisions b for all potential values of a . For each decision round, subjects were provided with a separate decision sheet to avoid changing prior decisions or copying information from previous choices. The decision sheets always indicated the partner's affiliation - signalling his or her country of origin - with whom a player was interacting at this moment. Having made all three decisions responders were asked to guess senders' actual transfers a and senders were asked to guess responders' back-transfers b for every potential value a . When making their transfer or back-transfer decisions, subjects did not know about the subsequent estimation task¹⁸.

Participants played the game against a counterpart from each subject pool without being informed about their payoffs before the very end of the entire experiment. Subjects took their decisions in complete anonymity. All participants were fully informed on all features of the experimental design and the procedures. Sessions lasted for about 90 minutes including reading the instructions. On average, subjects earned 23.9 US\$ (Germans 23.2 US\$, Israelis 24.6 US\$, and Palestinians 24.0 US\$).

4 Results

Our data analysis will proceed according to the research agenda formulated in section 2. As transfers and back-transfers in the trust game are typically regarded as measures of trust and trustworthiness we first analyze trust (senders' transfers) and responders' beliefs on trust. We then focus on trustworthiness (responders' back-transfers) and senders' beliefs on

¹⁷See Appendix B for the English translation of the original instruction. Instructions in German, Hebrew, and Arabic are available from the authors upon request.

¹⁸We decided on this timing of events because we wanted to avoid subjects' decisions being influenced by the subsequent belief statement.

trustworthiness. Finally, we study payoffs and performance. To facilitate reading we use abbreviations **G**, **I**, and **P** for identifying Germans, Israelis, and Palestinians. In general, we analyze the data according to the alphabetical order of the corresponding citizenship names. As we have no prior hypotheses concerning the direction of behavior, all statistical tests are two-sided.

4.1 The impact of culture on senders' trust and responders' trust beliefs

4.1.1 Trust standards

Our first research question is whether observed behavior matches findings of the literature on cross-cultural differences quoted in section 1. In particular, do trust (senders' transfers a) and beliefs on trust (responders' beliefs \tilde{a} on senders' transfers) differ across cultures and do different cultural standards exist with regard to trust? Moreover, do culture-specific standards also prevail in responders' beliefs on senders' trust? For this exploration we cross-culturally compare a and further contrast it with fellow-country responders' beliefs \tilde{a} .

	Population (I)	Measure (II)	To G (III)	To I (IV)	To P (V)	To all countries (VI)
(1)	From G	a	4.4 (3.3)	5.2 (3.5)	5.6 (3.3)	5.1 (3.3)
(2)		\tilde{a}	4.8 (3.1)	4.5 (3.5)	4.7 (3.3)	4.7 (3.2)
(3)	From I	a	3.9 (3.1)	3.9 (3.4)	2.9 (3.4)	3.6 (3.3)
(4)		\tilde{a}	4.8 (3.4)	4.2 (3.3)	3.7 (3.6)	4.2 (3.4)
(5)	From P	a	6.9 (1.6)	6.2 (2.5)	6.7 (2.3)	6.6 (2.1)
(6)		\tilde{a}	6.5 (2.5)	6.9 (2.7)	7.7 (2.7)	7.0 (2.6)
(7)	From all countries	a	5.1 (3.0)	5.1 (3.3)	5.1 (3.4)	5.1 (3.2)
(8)		\tilde{a}	5.4 (3.1)	5.2 (3.4)	5.4 (3.6)	5.3 (3.3)

Table 1: Senders' average transfers a and responders' average beliefs \tilde{a} on transfers a across countries in Experimental Currency Unit (ECU) (Example for reading the table: German senders transfer on average 4.4 to Germans, 5.2. to Israelis, and 5.6 to Palestinians. German responders believe to receive 4.8 from Germans, 4.5 from Israelis, and 4.7 from Palestinians).

Table 1 shows average a and \tilde{a} from and toward each of the three countries separately (columns III to V and lines 1 to 6). Column VI and lines 7 and 8 display aggregated transfers and beliefs from and toward all countries. Numbers in parentheses are standard deviations (SD).

Transfers a

According to predictions of standard game theory, we should observe $a^* = 0$. In line with the literature (see for an overview e.g., CARDENAS and CARPENTER, 2004a) we find participants in all subject pools to invest considerable amounts, however. Aggregated over all countries, 5.1 ECU are transferred to responders, with median = 5, mode = 5, and $SD = 3.2$. In **G**, in 86.7% of all possible cases senders transfer a positive amount, in **I** the percentage is 66.7%, and in **P** it reaches even 100%.

Full trust is displayed most frequently in **G** (17.8%), followed by **P** (13.3%) and **I** (6.7%); see Figure 1a. Hence, the standard game theoretic prediction is to be rejected for all countries and for all matchings. It is remarkable that subjects display a substantial level of trust in a totally anonymous inter-cultural context, in particular given the historical and political background of the participating countries like **G** and **I** or **I** and **P** and a complete anonymous interaction among strangers.

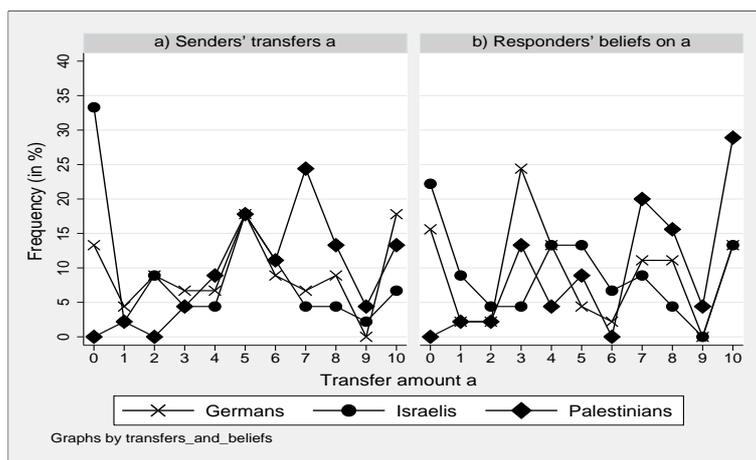


Figure 1a, b: a) Senders' transfers a b) Responders' beliefs \bar{a} per country.

We next answer our first research question, i.e. whether different trust standards predominate in the participating subject pools. To this end, we make a cross-cultural comparison of aggregated trust behavior by averaging transfers from a specific subject pool to the responders of all three countries (lines 1, 3, 5 and column VI in Table 1).

Result T1¹⁹: *Aggregated trust is different in all three countries.*

SUPPORT: Aggregated transfers are significantly different across countries ($p = .000$, Kruskal-Wallis test). Highest transfers are made by **P** (6.6 ECU), followed by **G** (5.1 ECU). Lowest trust is being shown by **I** (3.6 ECU) yielding the relation $a_P > a_G > a_I$. A pairwise cross-cultural analysis of individual transfers shows significant differences as well with $a_G > a_I$, $a_P > a_G$, and $a_P > a_I$ ($p = .040, p = .022, p = .000$, Mann-Whitney-U test)²⁰.

Result T1 confirms our conjecture that different levels of trust exist in the participating subject pools. This is in accordance with findings of INGLEHART et al. (2001) and MACKIE (2001) on differences in trust attitudes across and among cultures. Similar results are found in

¹⁹T refers to *Trust*.

²⁰A separated analysis by comparing the intra-cultural trust levels only gives a similar result. On average, **P** transfer 6.7 ECU to their fellow-countrymen, **G** transfer 4.4 ECU, and **I** 3.9 ECU.

bargaining experiments showing strong behavioral differences between subjects from different countries (e.g., HENRICH et al., 2001; BORNHORST, ICHINO, SCHLAG, and WINTER, 2007).

Beliefs \tilde{a}

In line with what senders actually did, responders expect a substantial level of trust behavior. Aggregated over all countries, responders believe that on average 5.3 ECU will be transferred, with median = 5, mode = 10, and $SD = 3.3$. In **G** and **I**, 86.7% of the responders believe in at least one positive transfer, in **P** the percentage is 100%. Full trust is most frequently expected by **P** (28.9%), followed by **G** and **I** (13.0% each); see Figure 1b. Responders in all countries and matchings do not expect the standard game theoretic prediction of zero-transfers. Note that responders hold positive expectations concerning trusting behavior also in the inter-cultural exchange.

In the next step we deal with responders' beliefs on senders' trust by comparing aggregated beliefs \tilde{a} (lines 2, 4, 6 and column VI in Table 1) across cultures as we did with transfers in the previous step of analysis.

Result T2: *Aggregated beliefs on senders' trust are different across countries.*

SUPPORT: Aggregated beliefs \tilde{a} are significantly different across countries ($p = .000$, Kruskal-Wallis test). Highest beliefs are stated by **P** (7.0 ECU), second highest by **G** (4.7 ECU), and lowest by **I** (4.2 ECU). Thus, the same relation as for transfers also holds for beliefs, namely $\tilde{a}_{\mathbf{P}} > \tilde{a}_{\mathbf{G}} > \tilde{a}_{\mathbf{I}}$. A pairwise cross-cultural analysis of individual beliefs shows significant differences, with $\tilde{a}_{\mathbf{P}} > \tilde{a}_{\mathbf{G}}$ ($p = .000$) and $\tilde{a}_{\mathbf{P}} > \tilde{a}_{\mathbf{I}}$ ($p = .000$). **G**- and **I**-beliefs do not differ significantly ($p = .600$), (all Mann-Whitney-U test)²¹.

Result T2 shows beliefs on trust to be different across countries as it was found with actual trusting behavior. Furthermore, the relations of these differences correspond to our findings on actual trusting behavior and support the conjecture that actual behavior and stated beliefs are linked on the intra-cultural level.

To go one step further we now investigate whether responders have correct beliefs on actual behavior of fellow-country senders, i.e., we analyze whether intra-cultural trust and trustworthiness standards are also reflected in the intra-cultural expectations concerning actual behavior (diagonal of Table 1). Recall that in all countries senders' transfers and responders' beliefs are elicited independently from different subjects in different experimental sessions.

²¹A separated analysis on the intra-cultural level delivers a similar result. On average, **P** expect a transfer of 7.7 ECU from their fellow-countrymen, **G** expect 4.8 ECU, and **I** 4.2 ECU.

Result T3: *Responders' beliefs match well fellow-country senders' actual behavior.*

SUPPORT: Our data show no differences between intra-cultural transfers and fellow-countrymen's beliefs on these transfers. **G(I)**-senders on average transfer 4.4(3.9) ECU to their fellow citizens. In contrast, **G(I)**-responders believe to receive 4.8(4.2) ECU from **G(I)**-senders. The differences between transfers and beliefs are not statistically significant (**G**: $p = .812$, **I**: $p = .910$, Mann-Whitney-U test). The same finding holds for **P**. They transfer 6.7 ECU and expect to gain 7.7 ECU (**P**: $p = .162$, Mann-Whitney-U test).

Result T3 indicates that responders in each country do not expect different trusting behavior than senders from their own country actually display. Given the fact that actual behavior and beliefs are elicited from different groups within our subject pools we conclude that responders' trust estimation is shaped by their culture affiliation.

To summarize our answers to our first research question, we find different and well-pronounced trust levels as well as different levels of beliefs on trusting behavior across subject pools (Results T1 and T2). In addition, the country-specific transfer standards are reflected in responders' beliefs on trusting behavior (Result T3).

4.1.2 Ethnocentrism and trust

Our second research question concerns the ethnocentric players' tendency to assess other cultures in terms of their own. If such behavior prevailed in our experiment, one countries' responders' beliefs on senders' behavior in all three assigned subject pools should not differ substantially; moreover, responders would have correct beliefs on actual behavior of fellow-country senders but they would be wrong in evaluating senders' behavior in the foreign subject pools.

Result T4: *Responders in G and P do not differentiate in beliefs. I-responders expect lower transfers from P- than from G-senders.*

SUPPORT: **G(P)**-responders expect on average 4.8, 4.5, and 4.7 (6.5, 6.9, and 7.7) ECU from their counterparts; see Table 1. **I**-responders on average expect lower transfers from **P**-senders (3.7 ECU) than from **G**- and **I**-senders (4.8 ECU and 4.2 ECU, respectively). A Friedman test shows that there is no statistical evidence that **G**- and **P**-responders do differentiate between subject pools (**G**: $p = .733$; **P**: $p = .665$). Contrary, **I**-responders do expect significantly higher transfers from **G**-senders than from **P**-senders ($p = .031$, Wilcoxon-signed-ranks test).

The above result shows practically no discrimination in beliefs among **G**, **I**, and **P**. Conducting a pairwise analysis we find **I**-responders expecting **P**-senders to back-transfer lower amounts compared to **G**-senders.

A further facet of ethnocentrism is that responders might attribute the standards of their own country also to foreign subject pools. As Result T3 shows responders did quite correctly anticipate actual behavior of fellow-country senders. To further investigate the influence of ethnocentric attitudes on actions and beliefs we compare senders' actual within-country transfers and the beliefs fellow-country responders hold toward foreign-country senders. To give an example for **G**-participants, we compare intra-cultural **G**-transfers [mean = 4.4 in row (1), column (III) in Table 1] with **G**-beliefs on **I**-transfers [mean = 4.5, row (2), column (IV)], and **P**-transfers to **G** [mean = 4.7, row (2), column (V) in Table 1] .

Result T5: *Responders attribute the transfer standards of their own population also to foreign subject pools.*

SUPPORT: **G**-senders on average transfer 4.4 ECU to **G**-responders while **G**-responders expect 4.5 ECU from **I**-senders and 4.7 ECU from **P**-senders. **I**-senders on average send 3.9 ECU to **I**-responders; **I**-responders believe **G**-senders to send 4.8 ECU and **P**-senders to send 3.7 ECU. **P**-senders on average transfer 6.7 ECU to **P**-responders while **P**-responders believe **G**-senders to transfer 6.5 ECU and **I**-senders to transfer 6.9 ECU. There is no evidence in the data that transfers and beliefs differ in the different matchings (**G** vs. **I**: $p = .977$; **G** vs. **P**: $p = .812$; **I** vs. **G**: $p = .486$, **I** vs. **P**: $p = .892$; **P** vs. **G**: $p = .926$, **P** vs. **I**: $p = .961$, Mann-Whitney-U test).

Result T5 shows that the standards which exist in a specific country/culture also seem to bias beliefs about foreigners' behavior. This is plausible when taking into account that subjects on average can form expectations only on what they know and on what they are familiar with. Cultural standards and norms are passed on from generation to generation. As SCHOTTER and SOPHER (2006), BOYD and RICHARDSON (1985), and CAVALLI-SFORZA and FELDMAN (1981) have argued and as FALK and ZEHNDER (2007) have shown, behavioral norms shape subjects' behavior which in turn forms their expectations.

As to our second research question subjects in our experiment seem to evaluate other cultures and their representatives' behavior in terms of their own. Result T4 shows that **G** and **P** do not differentiate in beliefs among assigned interaction partners, only **I** do so. Result T5 emphasizes this finding even more by showing that subjects do not only state similar beliefs toward foreigners but they also attribute the standards of their own culture to foreigners.

Our third research goal concerns actual discriminative behavior. If discrimination prevailed we should observe that responders are treated differently conditioned on their subject-pool affiliation. To answer our third research question we study senders' transfers a to each of the three subject pools separately (Table 1, lines 1, 3, 5 and columns III, IV, V).

Result T6: *I- and P-senders do not distinguish between subject pools. G-senders transfer less to fellow-country responders than to P-responders.*

SUPPORT: Table 1 shows that **G** on average transfer most to **P** (5.6 ECU), second most to **I** (5.2 ECU), and least to **G** (4.4 ECU), thereby discriminating compatriots against **P** ($p = .040$, Wilcoxon-signed-ranks test). **I**, on average, show similar trust to **G** and **I** (3.9 ECU). However, they transfer only 2.9 ECU to **P** - the smallest average transfer in our study (see Table 1). **P**, on the other hand, show the highest trusting behavior of all three countries transferring on average 6.9 ECU to **G**, 6.7 ECU to **P**, and 6.2 ECU to **I**. Although differences in averages are found, a Friedman test does not indicate any discrimination of **I**-senders against or in favor of one country ($p = .308$); the same holds for **P** ($p = .582$).

Interestingly, and contrary to what might be expected from the ethnocentrism literature, there is practically no discrimination among **G**, **I**, and **P**. However, we find some self-discrimination of **G** in trust behavior and **I**-responders expecting **P**-senders to back-transfer lower amounts than **G**-senders.

Our findings on trust decisions lead to the conclusion that **G**- and **I**-responders are likely to be surprised when interacting with **P**. **P**-senders significantly outperform **G(I)**-responders' expectations (4.7(3.7)) by transferring 6.9(6.2) ECU ($p = .055$, $p = .037$, Mann-Whitney-U test). Contrary, **P** might be disappointed by low **I**-transfers (2.9 ECU) which do not meet their expectations (6.9 ECU), ($p = .002$, Mann-Whitney-U test). These differences are remarkable. They might lead to a false perception of the intentions underlying the counterpart's decisions. As we have shown, there was no discrimination against foreign responders. Actions might be perceived as discriminative, however, or even hostile if subjects do not know the behavioral standards prevailing in the foreign culture. In a repeated game situation, this ignorance might easily lead to a breakdown of trust.

4.2 The impact of culture on reciprocity and senders' reciprocity beliefs

In this subsection, we investigate reciprocity behavior (responders' back-transfers b) and senders' beliefs \tilde{b} on responders' reciprocity. We base our analysis on the full response vectors gained by the strategy method. We report results on b (\tilde{b}) for *each* received amount $3a$, $a > 0$, with $3a = 3, 6, \dots, 27, 30$. See also Tables A1-A6 as well as Figures A1-A6 in the Appendix A and B.

4.2.1 Reciprocity standards

Our first research question with regard to reciprocity is whether we find different levels of b and \tilde{b} across cultures. Do different cultural standards exist with regard to reciprocity comparable to our findings on trust? Moreover, do culture-specific norms also prevail in senders' beliefs and expectations on responders' reciprocity? To answer these questions we first compare b across countries and then we compare b with fellow-country responders' beliefs \tilde{b} .

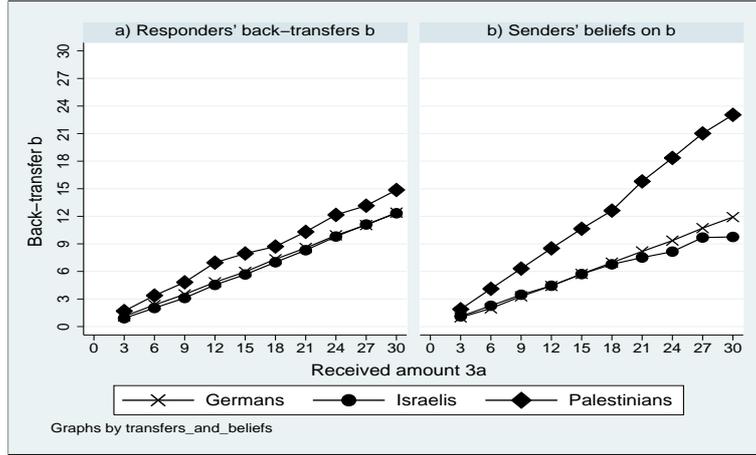


Figure 2a, b: a) Responders' average back-transfers b and b) senders' average beliefs \tilde{b} on back-transfers b across countries.

Figure 2 shows average b and \tilde{b} dependent on $3a$ for each of the three countries separately.

Back-transfers b

According to predictions of standard game theory, we should observe $b^* = 0$ for all values of $3a$. As it has been found in previous experiments so far, subjects in all three of our subject pools back-transfer considerable amounts. In 83% of all cases where $a > 0$, responders made a positive back-transfer. Aggregated over all countries and all received amounts $3a$, a fraction of 43.3% is back-transferred to senders, with median = 44.4%, mode = 0, and $SD = 29.0$. Among **G (I)**, responders back-transfer on average 1.11 to 12.40 ECU (0.91 to 12.33 ECU), the mean being 6.70 (6.48) ECU. **P** transfer 1.69 to 14.87 ECU back with an average of 8.40 ECU. This yields the relation $b_P > b_G > b_I$. All back-transfers b grow in absolute terms when transfers a increase (see also Figure 2a and Table A1).

Following our first research question, we investigate whether different standards of reciprocity exist in the participating subject pools. We, therefore, make a cross-cultural comparison of aggregated reciprocity behavior b for each received amount $3a$.

Result R1²²: *Aggregated back-transfers are different (in the lower range of received amounts $3a$) when **P**-responders are involved.*

SUPPORT: Aggregated back-transfers are (weakly) significantly different across countries in the lower range of received amounts, $3a = 3, 6, \dots, 15$ ²³. When **P** is involved, a pairwise cross-cultural analysis of individual back-transfers shows (highly) significant differences in the lower range of senders' tripled transfers with $b_{\mathbf{P}} > b_{\mathbf{G}} (b_{\mathbf{I}})$ ²⁴. The data show no evidence for differences between **G** and **I**²⁵.

Result R1 shows different levels of reciprocity to exist in interactions with **P**. In particular, we find differences for lower transfers a . Contrary, levels of reciprocity do not differ between **G** and **I**.

Beliefs \tilde{b}

We next analyze senders' beliefs \tilde{b} on responders' back-transfers b . We again use the complete dataset of response vectors gained by the strategy method.

In line with what responders actually did, senders expect substantial positive reciprocity behavior. Senders in all countries and matchings do not expect the standard game theoretic prediction of zero back-transfers. They correctly anticipate substantial high reciprocity behavior also in the inter-cultural context. In 88.7% of all cases where $a > 0$, senders anticipate a positive back-transfer. Aggregated over all countries and all received amounts $3a$, a fraction of 48.6% is believed to be back-transferred with median = 45.8%, mode = 33.3%, and $SD = 29.1\%$. In **G** (**I**), senders expect an average return of 1.00 to 11.91 ECU (1.11 to 9.76 ECU), the total mean being 6.35 (5.89) ECU. **P** expect a back-transfer of 1.89 to 23.04 ECU with a total mean of 12.23. Thus the same relation $\tilde{b}_{\mathbf{P}} > \tilde{b}_{\mathbf{G}} > \tilde{b}_{\mathbf{I}}$ holds for beliefs \tilde{b} as for back-transfers b . Beliefs on reciprocity also rise in absolute terms in $3a$ (see also Figure 2b and Table A2).

We next deal with our first research question concerning senders' beliefs on responders' back-transfers. As we did with back-transfers in the previous step of analysis we compare aggregated beliefs \tilde{b} across cultures contingent on the received amount $3a$.

Result R2: *Aggregated beliefs on responders' reciprocity are different when **P** is involved.*

²²R denotes Reciprocity.

²³Kruskal-Wallis test, all $p < .05$.

²⁴Mann-Whitney-U test, **P** vs. **G**: all $p < .05$; **P** vs. **I**: all $p < .05$.

²⁵A separated analysis by only comparing the intra-cultural back-transfer levels delivers a similar result. On average, **G** and **I** do not differ whereas **I** and **P** are different in the lower range.

SUPPORT: Aggregated beliefs \tilde{b} are highly significantly different across countries (for all transfers $3a$, $p = .000$, Kruskal-Wallis test). A pairwise cross-cultural analysis of individual beliefs on back-transfers shows (weakly) significant differences in the higher range of senders' tripled transfers ($3a = 24, 27, 30$) when comparing **G** and **I**²⁶. In interactions with **P**, differences are highly significant over the whole range of $3a$ (for all transfers $3a$, $p = .000$, Mann-Whitney-U test).

In line with our results on trust beliefs, Result R2 shows different levels of beliefs on reciprocity to exist in exchanges with **P**. We also find differences in the higher range of $3a$ between **G** and **I**.

In the following we investigate whether senders hold correct beliefs on reciprocity like responders did on trust. To this end, we compare senders' beliefs \tilde{b} with fellow-country responders' back-transfers b (Table A3). Recall that b and \tilde{b} are elicited independently from different experimental subjects in different sessions.

Result R3: *Senders' beliefs in **G** and **I** match well fellow-country responders' actual behavior. **P**-senders overestimate reciprocity behavior in their own country.*

SUPPORT: Comparing **G**-responders' back-transfers (1.13 to 12.53 ECU) and **G**-senders' beliefs (0.93 to 11.80 ECU), we find no statistical difference²⁷. The same holds for **I**-responders (0.93 to 13.33 ECU) and **I**-senders (1.20 to 10.00 ECU)²⁸ as well as for **P** in the lower range of $3a$ (3,..., 12)²⁹. In the upper range, however, **P**-senders expect (highly) significantly higher back-transfers (11.53 to 23.33 ECU) from fellow-country responders than these actually return (7.47 to 13.87 ECU)³⁰.

Result R3 shows that in **G** and **I** reciprocity norms seem to exist such that senders in both countries do not expect reciprocity behavior different from what responders in their own country actually display. This supports the conjecture that reciprocity behavior in **G** and **I** is also shaped by culture-specific standards. In **P**, this conclusion holds for the lower range of $3a$ only.

To summarize our findings with regard to our first research question, we find very similar reciprocity patterns as well as similar levels of beliefs on trusting behavior in **G** and **I**. **P**-participants, however, partly back-transfer higher amounts compared to **G** and **I** and hold

²⁶Mann-Whitney-U test, $p(24) = .032$, $p(27) = .054$, $p(30) = .017$

²⁷Mann-Whitney-U test, $p \geq .290$.

²⁸Mann-Whitney-U test, $p \geq .179$

²⁹ b : 1.60 to 6.40 ECU; \tilde{b} : 1.93 to 8.53 ECU; Mann-Whitney-U test, $p \geq .127$.

³⁰Mann-Whitney-U test, all $p < .05$.

significantly higher beliefs on back-transfers when compared to **G** and **I** (Results R1 and R2). In addition, the country-specific back-transfer standards are reflected in **G**- and **I**-senders' beliefs on trustworthiness. **P**-senders partly overestimate reciprocity behavior in their own country (Result T3).

4.2.2 Ethnocentrism and reciprocity

Our second research question is concerned with ethnocentric players' tendency to assess other cultures in terms of their own. If such behavior prevailed in our experiment, country's senders' beliefs on responders' behavior in all three assigned subject pools should not differ substantially. Moreover, senders should have correct beliefs on actual behavior of fellow-country responders but they should be wrong in evaluating responders' behavior in the foreign subject pools.

Result R4: *G*-senders do not differentiate in beliefs whereas *I* and *P* do.

SUPPORT: **G**-senders expect back-transfers from **G** between 0.93 and 11.80 ECU, from **I** between 1.00 and 11.53 ECU and from **P** between 1.07 to 12.40 ECU. They appear not to differentiate in their beliefs between countries as a pairwise comparison shows³¹. **I**-senders do distinguish according to senders' origin. They anticipate back-transfers from 1.40 to 10.93 ECU from **G**, from 1.20 to 10.00 ECU from **I**, and from 0.73 to 8.33 ECU from **P**. A pairwise comparison shows no discrimination between **G**- and **I**-responders³². Yet, **I**-senders believe to be returned lower amounts from **P** than from **G**³³ and **I**³⁴. **P**-senders also discriminate. They believe that **G**-responders will back-transfer from 2.00 to 24.07 ECU, that **I**-responders will return from 1.73 to 21.73 ECU, and that **P**-responders will pay back from 1.93 to 23.33 ECU. While they slightly discriminate **I** vs. **G**³⁵ and **P** vs. **I**³⁶ they slightly favor **G** vs. **P**³⁷ (see also Table A4 and Figures A4-A6).

The above results show no discrimination in beliefs for **G**-senders. Contrary, **I** and **P**-senders discriminate in their beliefs. **I**-senders expect lower back-transfers from **P**-responders. **P**-senders expect **G**-responders to back-transfer more compared to **P**- and **I**-responders. They expect **I**-responders to back-transfer less compared to **P**-responders.

³¹**G** vs. **I**: $p \geq .234$; **G** vs. **P**: $p \geq .133$; and **I** vs. **P**: $p \geq .172$, Wilcoxon-signed-ranks test.

³² $p \geq .375$, Wilcoxon-signed-ranks test.

³³ $3a = 3, \dots, 30$; Wilcoxon-signed-ranks test, $p(3) = .008$, $p(6) = .016$, $p(9) = .027$, $p(12) = .004$, $p(15) = .001$, $p(18) = .001$, $p(21) = .054$, $p(24) = .012$, $p(27) = .063$, $p(30) = .063$.

³⁴ $3a = 9, \dots, 27$; Wilcoxon-signed-ranks test, $p(9) = .008$, $p(12) = .023$, $p(15) = .023$, $p(18) = .008$, $p(21) = .010$, $p(24) = .023$, $p(27) = .055$.

³⁵ $3a = 9, 12, 18, 24$; Wilcoxon-signed-ranks test, $p(9) = .056$, $p(12) = .010$, $p(18) = .066$, $p(24) = .043$.

³⁶ $3a = 12$; $p = .043$, Wilcoxon-signed-ranks test.

³⁷ $3a = 6, \dots, 21$; Wilcoxon-signed-ranks test, $p(6) = .086$, $p(9) = .031$, $p(12) = .051$, $p(15) = .002$, $p(18) = .002$, $p(21) = .041$.

A further facet of ethnocentrism is that responders might attribute the standards of their own country also to foreign subject pools. As Result R3 shows **G**- and **I**-senders have quite correctly anticipated actual behavior of fellow-country responders. To further investigate the influence of ethnocentric attitudes on actions and beliefs we compare responders' actual within-country back-transfers with the beliefs fellow-country senders hold toward foreign-country responders. For **P**-participants for instance, we make an intra-cultural comparison of **P**-back-transfers with **P**-beliefs on **G**- and **I**-back-transfers.

Result R5: ***G**- and **I**-senders attribute the standards of their own country also to the other two foreign subject pools. **P** expect higher back-transfers from foreign than from fellow-country responders, mainly in the upper range of 3a.*

SUPPORT: **G**-senders assign the standards of their own country to the two foreign subject pools. They believe that **I**-responders return from 1.07 to 12.40 ECU and that **P**-responders back-transfer between 1.00 and 11.53 ECU. **G**-responders assert to back-transfer from 1.13 to 12.53 ECU to fellow-country senders. There is no evidence in the data that differences exist³⁸. **I**-senders also appear ethnocentric by believing that **G**-responders return from 1.40 to 10.93 ECU and **P**-responders back-transfer between 0.73 and 8.33 ECU. In fact, **I**-responders state to back-transfer from 0.93 to 13.33 ECU to fellow-country senders. There is also (nearly) no evidence that **I** distinguish between subject pools³⁹. The standards **P**-senders attribute to foreign subject pools are different from actual fellow-country responders' behavior. They expect **G**-responders to return from 2.00 to 24.07 ECU and **I**-responders to back-transfer between 1.73 and 21.73 ECU. In fact, **P**-responders are willing to back-transfer to fellow-country senders from 1.60 to 13.87 ECU only. Thus, **P**-senders expect significantly higher paybacks from **G**⁴⁰- and from **I**-responders⁴¹ than **P**-responders actually return, particularly in the upper range of 3a (see also Table A5).

Result R5 shows that **G**- and **I**-senders behave ethnocentrically with regard to reciprocity. This holds true although we do not find noticeable differences in levels of reciprocity and reciprocity beliefs. Similar to trust behavior, this is plausible when taking into account that subjects on average can only form correct beliefs on what they know and on what they are familiar with (see FALK and ZEHNDER, 2007). **P**-senders, however, do not only overestimate reciprocity behavior of their fellow-country responders. They also expect much higher returns from foreign responders than responders from their own subject pool are actually willing to

³⁸Mann-Whitney-U test; **I**: $p \geq .400$; **P**: $p \geq .654$.

³⁹Mann-Whitney-U test; **G**: $p \geq .194$; **P**: $p \geq .091$.

⁴⁰3a = 6,..., 30; Mann-Whitney-U test, $p(6) = .083$, $p(9) = .062$, $p(12) = .007$, $p(15) = .036$, $p(18) = .004$, $p(21) = .010$, $p(24) = .003$, $p(27) = .009$, $p(30) = .011$.

⁴¹3a = 18,..., 30; Mann-Whitney-U test, $p(18) = .035$, $p(21) = .019$, $p(24) = .019$, $p(27) = .022$, $p(30) = .065$.

back-transfer.

Our third research goal with regard to reciprocity deals with actual discriminative behavior. If discrimination prevailed, we should observe that senders are treated differently dependent on their subject-pool affiliation. To answer our research question three we study back-transfers b and senders' beliefs \tilde{b} on back-transfers to each of the three subject pools separately.

Result R6: *Responders in the three subject pools do not discriminate substantially.*

SUPPORT: Table A6 shows that **G**-responders back-transfer from 1.13 to 12.53 ECU to **G**, from 1.07 to 12.07 ECU to **I**, and from 1.13 to 12.60 ECU to **P**. We do not find evidence in the data that they differentiate between countries. A pairwise comparison confirms this result⁴². **I**-responders back-transfer from 0.93 to 12.00 ECU to **G**, from 1.06 to 13.81 ECU to **I**, and from 0.87 to 11.67 ECU to **P**. They also do not distinguish with regard to senders' origin as a pairwise comparison shows⁴³. **P**-responders back-transfer from 1.73 to 16.93 ECU to **G**, from 1.73 to 13.80 ECU to **I**, and from 1.60 to 13.87 ECU to **P**. While they favor **G** compared to **P** in the upper range of $3a$ ⁴⁴ we find (nearly) no evidence in the data that **I**- and **P**-senders are treated differently⁴⁵ (see also Figures A1-A3).

Again, contrary to our expectation based on the ethnocentrism literature, the above results show practically no discrimination among **G**, **I**, and **P**. We find some positive discrimination in favor of **G**-senders by **P**-responders, though.

Our findings on reciprocity lead to the conclusion that **G**- and **I**-responders most likely will be surprised when interacting with **P** as **P**-responders outperform **G**(**I**)-senders' expectations. Contrary, **P** will be disappointed by low **G**- and **I**-back-transfers which do not meet their expectations. As we have shown, there was no substantial discrimination against foreign senders. **P**, not knowing the behavioral standards prevailing in **G** and **I**, might perceive **G** and **I**-actions as discriminative.

Our results on reciprocity are not as clear-cut as those on trust. Two points lend themselves as possible explanations. For one thing, responders (senders) had to make not only one single decision (state one belief) but had to state a (belief on a) back-transfer for each possible outcome of $3a$. It is not implausible that not every participant makes consistent decisions

⁴²**G** vs. **I**: $p \geq .625$; **G** vs. **P**: $p \geq .813$; and **I** vs. **P**: $p \geq .625$, Wilcoxon-signed-ranks test.

⁴³**I** vs. **G**: $p \geq .188$; **G** vs. **P**: $p \geq .625$; and **I** vs. **P**: $p \geq .125$, Wilcoxon-signed-ranks test.

⁴⁴ $3a = 12, 24, \dots, 30$; Wilcoxon-signed-ranks test, **G** vs. **P**: $p(12) = .039, p(24) = .015, p(27) = .047, p(30) = .031$.

⁴⁵**G** vs. **I**: $p \geq .082$; **I** vs. **P**: $p \geq .148$, Wilcoxon-signed-ranks test.

in such a multi-statement task. As a result, a higher variance might blur existing cross-cultural differences. Second, **P**-participants behave differently when deciding on reciprocity as compared to decisions on trust. In particular, **P**-responders state higher back-transfers than responders in the other two subject pools and, more importantly, **P**-senders highly overestimate responders' back-transfers not only from foreign countries but also from their own subject pool. This leads to significant differences in interactions with **P**-participants as for instance with reciprocity norms, ethnocentrism, and discrimination. We will discuss this finding in section 5.

4.3 Return on investment

In this subsection, we analyze the monetary consequences of trust and reciprocity behavior. To this end, we study participants' performance, i.e., senders' and responders' *actual* payoffs not including the show-up fee. The actual payoff per round of the sender is calculated by his initial endowments $X = 10$ minus his transfers a plus the corresponding responders' actual back-transfers b^{act} ⁴⁶. Likewise, the actual payoff of the responder is given by her initial endowments $X = 10$ plus the received (tripled) transfers a from the senders minus the corresponding actual back-transfers b^{act} to the senders⁴⁷.

On average, responders receive highly significantly higher payoffs (30.4 US\$) than senders (17.4 US\$), ($p = .000$, Mann-Whitney-U test). This is due to the fact that on average responders do not share the *total* pie ($X - a + X + 3a$) equally but instead back-transfer $b^{act} < 2a$, i.e., less than two thirds of $3a$. When aggregating senders' and responders' payoffs, **I** gain the highest amount (24.6 US\$), followed by **P** (24.0 US\$) and **G** (23.2 US\$). Payoff levels do not differ across countries ($p = .900$, Kruskal-Wallis test).

Does cross-border trust pay in the trilateral trust game? Highest b^{act} are made by **P** (53.1%), followed by **G** (41.0%) and **I** (39.0%) (all b^{act} as a percentage of a). As with b , this yields the relation $b_{\mathbf{P}}^{act} > b_{\mathbf{G}}^{act} > b_{\mathbf{I}}^{act}$. On average, *absolute* amounts of back-transfers increase in a in all matchings, implying that the more senders trust the more they are rewarded in absolute terms. For *relative* rewarding the picture is different. While **G** and **I** reward higher transfers a with higher back-transfers b^{act} , **P** do the opposite. The larger the amount $3a$ **P**-responders receive, the smaller is their back-transfer b^{act} . Moreover, reciprocity levels in **G** and **I** are lower than in **P** even though $b_{\mathbf{G}}^{act}$ and $b_{\mathbf{P}}^{act}$ increase over a . Discrimination in reciprocity against or in favor of a specific country cannot be found (see also Figures A1-A3).

$$^{46} \pi_{sender} = 3X - \sum_{i=1}^3 a + \sum_{i=1}^3 b^{act}.$$

$$^{47} \pi_{responder} = 3X + \sum_{i=1}^3 a - \sum_{i=1}^3 b^{act}.$$

Finally, we calculate the average rate of return $R = 100 \cdot \left(\frac{b^{act}}{a} - 1 \right) - 100$, i.e., the percentage a sender receives back in addition to his investment. R indicates how well trust pays off (see Table 5).

Population	To G	To I	To P	To all countries
From G	24.1	16.9	73.6	40.0
From I	17.0	30.0	62.3	35.0
From P	23.8	12.8	44.5	27.0
From all countries	22.0	18.8	59.4	33.4

Table 5: Average rates of return R in %.

Over all matchings, trusting **P**-responders yields the highest return R for all senders. This holds in particular for **G**-senders (**I**-senders) getting 73.6% (62.3%) back in addition to their transfers to **P** (see Table 5). Trusting **I**, however, is not as profitable for **P**-senders (**G**-senders) as trusting people from the other two countries. In interaction with **I** they realize the lowest (second lowest) return R of 12.8% (16.9%). Additionally, for **I**-senders, it turns out to be a bad bargain to trust **G** because of getting a return R of only 17.0%.

5 Discussion

We have investigated actions and beliefs in an inter-cultural trilateral trust game played among Germans, Israelis, and Palestinians. Following the work of SUMMER (1906), ADAMS (1951), SCHOPMEYER and FISHER (1993), and KALIN and BERRY (1994), we studied the influence of ethnocentrism on trusting and reciprocity behavior.

In a first step we explored whether different cultural standards of trust and trustworthiness exist within our subject pools. For trust behavior we find different and well-pronounced trust levels as well as different levels of beliefs on trusting behavior across countries. Furthermore, the country-specific transfer standards are correctly reflected in responders' beliefs on trusting behavior. For reciprocity, we find very similar patterns as well as analogous levels of beliefs on reciprocity for Germans and Israelis. In addition, in these populations the country-specific back-transfer standards are also reflected in beliefs on reciprocity behavior. Palestinians, however, back-transfer higher amounts and hold significantly higher beliefs on back-transfers when compared to Germans and Israelis. Moreover, Palestinian senders slightly overestimate reciprocity behavior in their own country.

How can we explain different levels of trust across our samples? One possible explanation for the low amount of trust found for Israelis is that the Israeli society suffers from a clearly significant and heated segmentation between secular and Orthodox Jews and between Arabs

and Jews⁴⁸. In such a segmented society, where a persistent economic and educational gap between the segments exists and continues to prevail among second-generation immigrants, comprehensive networks of trust are difficult to establish (see also FERSHTMAN and GNEEZY, 2001). Systematic mistrust toward fellow citizens of other origin inhibit the formation of such social capital. In Palestine we find a comparatively high level of trust. This is also reflected in high beliefs about corresponding back-transfers. Even though these beliefs are too optimistic (and therefore apparently not biased by ethnocentric considerations), the variance in Palestinian trusting behavior is significantly lower compared to Germans and Israelis⁴⁹. A further explanation is given by the so-called “cushion-hypothesis” (HSEE and WEBER, 1999) saying that members of socially collectivist cultures like the Arabian culture⁵⁰ can afford to take greater financial risks since their social networks insure them against catastrophic outcomes. For the same reason they have to be much more concerned about social risks like, e.g., losing face. Therefore, it is plausible that Palestinians acted in a socially desirable way to keep honesty and reputation and to not lose face. These are important values in Arabic societies.

Our second research question was related to whether our subjects ethnocentrically assign their home standards and norms to foreign cultures. We conjectured that subjects form correct beliefs on actual behavior of fellow-country counterparts but that they are wrong in guessing what their counterparts in the foreign subject pools actually do. Moreover, we expected that within a country beliefs about the behavior in all three assigned subject pools do not differ substantially. Our trust results show that the standards which exist in a specific society also seem to bias beliefs about foreigners’ behavior. This finding gets further support by the fact that our subjects show practically no discrimination in beliefs. For reciprocity the picture is not as clear-cut. Again, Germans and Israelis behave ethnocentrically with regard to reciprocity. This holds true although we do not find noticeable differences in the levels of reciprocity and reciprocity beliefs among these cultures. Furthermore, German senders show no discrimination in their beliefs whereas Israelis expect lower back-transfers from Palestinian responders. Moreover, Palestinians not only overestimate reciprocity behavior of their fellow-country responders. They also expect much higher returns from foreign responders than responders from their own society actually back-transfer. In addition, Palestinian senders

⁴⁸The two major ethnic groups are Ashkenazic Jews (European and American immigrants and their Israeli-born offspring) and Eastern Jews (Asian and African immigrants and their Israeli-born offspring).

⁴⁹ SD_G vs. $SD_P = 3.3$ vs. 2.1 ($p = .045$), SD_I vs. $SD_P = 3.3$ vs. 2.1 ($p = .037$), all exact two-sample permutation test for differences in variances.

⁵⁰The lowest score for the Arab World on any Hofstede-dimension is the Individualism score at rank 38, compared to a world average rank of 64 (source: www.geert-hofstede.com, see also Hofstede (2001)). “Collectivism” and “Individualism” are the opposite extremes of the same scale. Specific data for Palestinians are not available.

discriminate in their beliefs and expect German responders to back-transfer more compared to Palestinians and they expect Palestinians to back-transfer more compared to Israelis.

Why does ethnocentrism apparently have a significant impact on trusting decisions and why does the link between ethnocentrism and reciprocity seem less clear? The answer to this question may be found in the different characteristics of the two player-roles with regard to the perception of and the disposition towards risk. According to DOUGLAS and WILDAVSKY (1983) and WEBER and HSEE (1998), different social groups fear different risks, perceive risks differently, and have different attitudes toward risk. The truster in our experiment has to take a decision that is associated with the risk of a possible loss. Having no information on the responder's actions but possibly taking into account that the trustee's decision might or might not be influenced by the sender's trust, the sender can only form a belief on how the responder might behave. The sender's very ambiguous situation might lead to higher levels of reasoning (NAGEL, 1995; BOSCH-DOMÈNECH, GARCÍA-MONTALVO, NAGEL, and SATORRA, 2002). Moreover, it may trigger culture-specific subconscious cognitive processes making actions more likely that are based on ethnocentric beliefs. The responder, on the other hand, plays a dictator game. As she faces less potential risk⁵¹, the above mentioned cognitive processes may play only a secondary role and therefore might have a lower impact on responders' decisions.

In a third step we analyzed another feature of ethnocentric behavior, namely whether one's own cultural norms and values tend to be considered superior to all others. Such attitudes would lead to discrimination against foreign players compared to fellow-country men. Contrary to our expectation, we found practically no discrimination among Germans, Israelis, and Palestinians in trusting behavior. Our trustworthiness data, however, show some self-discrimination of Germans and Israeli responders expecting Palestinians to back-transfer lower amounts than German senders. We observe practically no discrimination among Germans, Israelis, and Palestinians although we find some positive discrimination in favor of Germans by Palestinian responders.

Finding only small discrimination effects suggests that prejudices about foreign counterparts as well as taste-based discrimination do only play a minor role in our experiment. In fact, our data suggest the decision maker's cultural background to bias beliefs and wrong stereotypes. Cultural affiliation seems to shape beliefs which guide actions that in turn enforce beliefs, and so on.

⁵¹The responder, however, faces some degree of uncertainty about what the sender *actually* does.

Our last research question is related to the monetary outcome of our trilateral trust game. We find that on average responders receive higher payoffs than senders in all cultures. Israelis gain highest and Germans earn lowest amounts. Payoffs are not significantly different across countries, though. Trusting in Palestinians yields the highest return for all senders. Contrary, transfers to Israelis seems to be a bad deal for foreigners as they realize by far the lowest returns.

As ethnocentric participants attribute the standards of their own country also to foreign subject pools, direct interactions among strangers might result in substantial surprises and disappointments because of the cross-cultural differences. Yet, the underlying motive is *not* discrimination. Instead, the behavioral standards prevailing in the participating countries might create divergent beliefs about the other side's intentions. A false perception of negative (positive) discrimination might arise if subjects are ignorant about the divergence in trust and reciprocity between countries. It has been shown in the literature that a mismatch of decisions and expectations can lead to aggressive behavior e.g. in sequential interactions (BOSMAN and VAN WINDEN, 2002; BOSMAN, HENNIG-SCHMIDT, and VAN WINDEN, 2006). Thus, the clash of actual behavior and beliefs in inter-cultural interactions can trigger and enforce distrusting behavior.

6 Conclusion

Our study reveals important features of cross-cultural interactions in the trilateral context including the lack of main inter-cultural discrimination effects but the existence of different social standards and norms across countries. Our results support the view that conflicts may not only be triggered by discrimination but may also be enforced by the different social standards in the three societies. Our findings improve the understanding of cross-cultural interactions in quite a substantial way. First, culture-inherent standards and norms appear to affect trust and trustworthiness substantially more than stereotypes and prejudices. The belief-building process seems to be shaped and biased by own-populations standards and norms. Second, in our experiment, subjects interacted only once. Consequently, they could not react, e.g., to a perceived discrimination that in fact was a culturally biased lower level of trust and/or reciprocity. In repeated interactions, however, a misinterpretation of foreign-culture behavior as discriminative may easily lead to retaliation and may raise or intensify inter-cultural confrontations. Thus, the clash of cultures in repeated interactions bears huge potential for escalating conflicts.

The understanding of culture-specific or country-specific social mechanisms and their origins

leading to the observed behavior is crucial not only for the analysis of scenarios like the Israeli-Palestinian conflict. It is also important for the spread and use of scientific concepts and theories (e.g., management theories). Each culture has a unique and strong background with distinct characteristics that differentiate it from others and encourage various meanings of behavior that work for each. Observed behavior in a specific situation cannot be isolated from other processes taking place in a society because it interacts with what happens in the family, at school, in politics, and government. It is obviously also related to religion and beliefs about science itself. Therefore, behavioral standards and theories accepted in one culture might not be applicable globally and shouldn't be assumed as ideal for all cultures. (HOFSTEDE, 1993).

Based on our findings regarding potential sources of conflict, a key issue for future research is to develop a better understanding on how inter-cultural conflicts escalate and how they can be reduced in repeated scenarios. An intention-based approach (see e.g. MCCABE, RIGDON, and SMITH, 2003; FALK and FISCHBACHER, 2006) might be an appealing starting point. If for instance responders are informed about foreign-country senders' intentions/different cultural norms - Israeli (Palestinian) senders in general transfer low (high) amounts - they might equalize their beliefs and interpret sender's actions less ethnocentrically. This might lead to adapted expectations and reduce the danger of conflicts in repeated interactions.

References

- ADAMS, R.N. (1952): "Ethnocentrism and Ingroup Consciousness", *American Anthropologist*, Vol. 53(4), 598-600.
- AIGNER, D.J. and G.G. CAIN (1977): "Statistical Theories of Discrimination in Labor Markets", *Industrial and Labor Relations Review*, 30(2), 175-187.
- ARROW, K.J. (1971): "The Theory of Discrimination", *Working Papers 403*, Princeton University, Department of Economics, Industrial Relations Section.
- ARROW, K.J. (1972): "Gifts and Exchanges", *Philosophy and Public Affairs*, 1(4), 343-362.
- BECKER, G.S. (1957): "The Economics of Discrimination", University Of Chicago Press, 2. revised edition (1971).
- BELLEMARE, C. and S. KRÖGER (2007): "On Representative Social Capital", *European Economic Review*, 51(1), 183-202.
- BEN-NER, A. and L. PUTTERMAN (2001): "Trusting and Trustworthiness", *Boston University Law Review*, 81, 523-551.

- BERG, J., J. DICKHAUT, and K. MCCABE (1995): "Trust, Reciprocity, and Social History", *Games and Economic Behavior*, 10(1), 122-142.
- BOARINI, R., J.-F. LASLIER, and S. ROBIN (2002): "Interpersonal Comparison of Utility in Bargaining: Evidence from a Transcontinental Ultimatum Game", *Working Paper, Laboratoire d'Econométrie de l'Ecole Polytechnique, Paris*.
- BOND, M.H. and K. HWANG (1986): "The Psychology of the Chinese People", in: M. H. Bond (Ed.), *The Social Psychology of the Chinese People*, Hong Kong, Oxford University Press 1986, pp. 213-266.
- BORNHORST, F., A. ICHINO, K. SCHLAG, and E. WINTER (2007): "Trust and Trustworthiness among Europeans South - North Comparison", *CEPR Discussion Paper No. 4378*.
- BOSMAN, R. and F. VAN WINDEN (2002): "Emotional Hazard in a Power-to-Take Game Experiment", *The Economic Journal*, 112, 146-169.
- BOSMAN, R., H. HENNIG-SCHMIDT, and F. VAN WINDEN (2006): "Exploring Group Decision Making in a Power-to-Take Experiment", *Experimental Economics*, 9, 35-51.
- BOYD, R. and P. J. RICHARDSON (1985): "Culture and the Evolutionary Process", *Chicago University Press, Chicago*.
- BRANDTS, J. and G. CHARNESS (2000): "Hot vs. Cold Sequential Responses and Preference Stability in Experimental Games", *Experimental Economics*, 2(3), 227-238.
- BRISLIN, R.W. (1970): "Back-Translation for Cross-Cultural Research", *Journal of Cross-cultural Psychology*, 1, 185-216.
- BUCHAN, N.R., R.T.A. CROSON, and E.J. JOHNSON (2004a): "When Do Fairness Beliefs Influence Bargaining Behavior: Experimental Evidence from Japan and the United States", *Journal of Consumer Research*, 31(2004), 181-190.
- BUCHAN, N.R., R.T.A. CROSON, and E.J. JOHNSON (2004b): "The Boundaries of Trust: Own and Other's Actions in the US and China", *Journal of Economic Behavior and Organization*, 55(4), 485-504.
- CARDENAS, J.C. and J. CARPENTER (2004): "Experimental Development Economics: A Review of the Literature and Ideas for Future Research", *Middlebury College Working Paper, Middlebury College*.
- CAVALLI-SFORZA, L. L. and M. W. FELDMANN (1981): "Cultural Transmission and Evolution: A Quantitative Approach", *Princeton University Press, Princeton, NJ*.
- CASTRO, M.F. (2007): "Where are you from? Cultural Differences in Public Good Experiments", *Working paper, Department of Economics, Royal Holloway College, University of London, UK*.

- COOK, K.S. and R.M. COOPER (2003): "Experimental Studies of Cooperation, Trust, and Social Exchange", in: E. Ostrom and J. Walker (Eds.), *Trust and Reciprocity*, Russell Sage, New York (2003), pp. 209-244.
- CHUAH, S.H., R. HOFFMANN, M. JONES, and G. WILLIAMS (2007): "Do Cultures Clash? Evidence from Cross-National Ultimatum Game Experiments", *Journal of Economic Behavior and Organization*, 64(1), 35-48.
- CROSON, R.T.A. and N.R. BUCHAN (1999): "Gender and Culture: International Experimental Evidence from Trust Games", *The American Economic Review*, 89(2), 386-391.
- DOUGLAS, M. and A.B. WILDAVSKY (1983): "Risk and Culture: An Essay on the Selection of Technical and Environmental Dangers", *Journal for the Scientific Study of Religion*, Vol. 22, No. 2, 188-189.
- ECO, U. (2000): "Experiences in Translation", *Toronto Italian Studies*, University of Toronto Press, 2000.
- ENCYCLOPAEDIA BRITANNICA 2006: Retrieved January 11, 2008, from Britannica Concise Encyclopedia: <http://concise.britannica.com/ebc/article-9363955>.
- FALK, A. and U. FISCHBACHER (2006): "A Theory on Reciprocity", *Games and Economic Behavior*, Volume 54, Issue 2, 293-315.
- FALK, A. and C. ZEHNDER (2007): "Discrimination and In-Group Favoritism in a Citywide Trust Experiment", Zurich IEER Working Paper No. 318.
- FISCHER, R. (2006): "Congruence and Functions of Personal and Cultural Values: Do my Values Reflect my Culture's Values?", *Personality and Social Psychology Bulletin*, in press.
- FUKUYAMA, F. (1995): "Trust: The Social Virtues and the Creation of Prosperity", *The Free Press, New York*.
- FUKUYAMA, F. (2001): "Social Capital, Civil Society and Development", *Third World Quarterly*, 22(1), 7-20.
- FERSHTMAN, C. and U. GNEEZY (2001): "Discrimination in a Segmented Society: An Experimental Approach", *Quarterly Journal of Economics*, Feb. 2001, 351-377.
- GOERG, S., W. GÜTH, H. HENNIG-SCHMIDT, G. WALKOWITZ, and T. WEILAND (2008): "Distributive fairness in an intercultural ultimatum game", *Jena Economic Research Paper*, 2008-028.
- GREINER, B. (2004): "An Online Recruitment System for Economic Experiments", in: V. M. Kurt Kremer (Ed.), *Forschung und Wissenschaftliches Rechnen 2003*, GWDG Bericht 63, Gesellschaft für Wissenschaftliche Datenverarbeitung, Göttingen, pp. 79-93.

- GÜTH, W., C. SCHMIDT, and M. SUTTER (2003): "Fairness in the Mail and Opportunism in the Internet: A Newspaper Experiment on Ultimatum Bargaining", *German Economic Review*, 4(2), 243-265.
- HENRICH, J. (2000): "Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining among the Machiguenga", *American Economic Review*, 90(4), 973-979.
- HENRICH, J., R. BOYD, S. BOWLES, C. CAMERER, E. FEHR, H. GINTIS, and R. MCELREATH (2001): "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies", *American Economic Review*, 91(2), 73-78.
- HERRMANN, B., C. THÖNI, and S. GÄCHTER (2008): "Antisocial Punishment Across Societies Cooperation and Punishment in a Cross-Cultural Perspective", *Science* 319, 1362-1367.
- HOFSTEDE, G. (1993): "Cultural constraints in management theories", *Academy of Management Executive*, 7, 1, 81-94.
- HOFSTEDE, G. (2001): "Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations", 2nd Edition, *Sage Publications, Thousand Oaks*.
- HOLM, H. and A. DANIELSON (2005): "Tropic Trust versus Nordic Trust: Experimental Evidence from Tanzania and Sweden", *Economic Journal*, Vol. 115, No. 503, 505-532.
- HSEE, C.K. and E.U. WEBER (1999): "Cross-national differences in risk preferences and lay predictions", *Journal of Behavioral Decision Making*, 12, 165-179.
- INGLEHART, R., M. BASAÑEZ, and A. MORENO (2001): "Human Values and Beliefs", *University of Michigan Press, MI*.
- JACOBSEN, E. and A. SADRIEH (1996): "Experimental Proof for the Motivational Importance of Reciprocity", *Discussion Paper Serie B 386*, University of Bonn, Germany.
- KAHAN, D.M. (2008): "Cultural Cognition as a Conception of the Cultural Theory of Risk", *Cultural Cognition Project Working Paper No. 45*, Yale Law School.
- KALIN, R. and J.W. BERRY (1994): "Ethnic and multicultural attitudes", in: J.W. Berry and J.A. Laponce (eds.): *Ethnicity and culture in Canada: The research landscape*, Toronto, Canada: The University of Toronto Press.
- KNACK, S. and P. KEEFER (1997): "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation", *Quarterly Journal of Economics*, 112, 1251-1288.
- KOFORD, K. (2001): "Trust and Reciprocity in Bulgaria: A Replication of Berg, Dickhaut and McCabe", *Working paper, University of Delaware*.
- LAZZARINI, S., M. REGINA, A. RINALDO, and J.O. SIQUEIRA (2004): "Measuring Trust: An Experiment in Brazil", *Working Paper*.

- LUHMANN, N. (1979): "Trust: A Mechanism for the Reduction of Social Complexity", in: N. Luhmann (Ed.), *Trust and Power*, Wiley, New York (1979), pp. 4-103.
- MACKIE, G. (2001): "Patterns of Social Trust in Western Europe and their Genesis", in: K. Cook (Ed.), *Trust in Society*, Russell Sage Foundation, New York, pp. 245-282.
- MCCABE, K.A., M.L. RIGDON, and V.L. SMITH (2003): "Positive reciprocity and intentions in trust games", *Journal of Economic Behavior and Organization*, Vol. 52, 267-275.
- NAGEL, R. (1995): "Unraveling in Guessing Games: An Experimental Study", *American Economic Review*, 85(5), 1313-26.
- PARSONS, T. (1957): "The Social System", *New York, Free Press*.
- PHELPS, E.S. (1972): "The Statistical Theory of Racism and Sexism", *The American Economic Review*, 62(4), 659-661.
- PUTNAM, R.D., R. LEONARDI, and R.Y. NANETTI (1994): "Making Democracy Work: Civic Traditions in Modern Italy", Princeton, NJ: Princeton University Press, 1993.
- ROTH, A.E., V. PRASNIKAR, M. OKUNO-FUJIWARA, and S. ZAMIR (1991): "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study", *American Economic Review*, 81(5), 1068-1095.
- SCHOPMEYER, K.D., and B.J. FISHER (1993): "Insiders and Outsiders: Exploring Ethnocentrism and Cultural Relativity in Sociology Courses", *Teaching Sociology*, Vol. 21(2), 148-153.
- SCHOTTER, A., and B. SOPHER (2006): "Trust and Trustworthiness in Intergenerational Games: An Experimental Study of Intergenerational Advice", *Experimental Economics*, Vol. 9(2), 123-145.
- SCHWARTZ, S.H. (1992): "Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 countries, in: M. P. Zanna (Ed.), *Advances in Experimental Social Psychology*, 25, San Diego, CA, Academic Press, pp. 1-65.
- SELTEN, R. (1967): "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperimentes, in: H. Sauermann (Ed.), *Beiträge zur experimentellen Wirtschaftsforschung*, J.C.B. Mohr (Paul Siebeck), Tübingen, pp. 136-168.
- SUMMER, W.G. (1906): "Folkways", Boston, Ginn, 1906.
- TAN, J.H.W. and C. VOGEL (2005): "Religion and Trust: An Experimental Study, *Department of Economics, European University Viadrina Frankfurt (Oder), Germany*, Discussion Paper No. 240.
- THE COLUMBIA ENCYCLOPEDIA, 6th ed. New York: Columbia University Press, 2001-07.

- WALKOWITZ, G., C. OBERHAMMER, and H. HENNIG-SCHMIDT (2004): “Experimenting Over a Long Distance - A Method to Facilitate Inter-cultural Experiments, *Working Paper, Department of Economics, University of Bonn*.”
- WEBER, E.U. and C.K. HSEE (1998): “Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes towards Perceived Risk”, *Management Science*, Vol. 44, Issue 9, 1205-1217.
- WILLIAMSON, O.E. (1981): “The Economics of Organization: The Transaction Cost Approach, *The American Journal of Sociology*, 87(3), 548-577.
- WILLINGER, M., C. KESER, C. LOHMANN, and J.-C. USUNIER (2003): “A Comparison of Trust and Reciprocity between France and Germany: Experimental Investigation Based on the Investment Game, *Journal of Economic Psychology*, 24(3), 447-466.

Appendix A: Tables

Received $3a$	3	6	9	12	15	18	21	24	27	30	AV
Germany	1.11	2.36	3.51	4.82	5.96	7.33	8.56	9.93	11.04	12.40	6.70
Israel	0.91	2.02	3.11	4.53	5.67	7.00	8.31	9.80	11.11	12.33	6.48
Palestine	1.69	3.38	4.82	6.96	7.96	8.71	10.31	12.16	13.16	14.87	8.40
AV	1.24	2.59	3.81	5.44	6.53	7.68	9.06	10.63	11.77	13.20	7.19

Table A1: Responders' aggregated back-transfers b .

Received $3a$	3	6	9	12	15	18	21	24	27	30	AV
Germany	1.00	1.98	3.29	4.42	5.73	6.96	8.18	9.36	10.71	11.91	6.35
Israel	1.11	2.27	3.47	4.44	5.71	6.78	7.51	8.16	9.69	9.76	5.89
Palestine	1.89	4.11	6.31	8.51	10.64	12.62	15.80	18.36	21.02	23.04	12.23
AV	1.33	2.79	4.36	5.79	7.36	8.79	10.50	11.96	13.81	14.90	8.16

Table A2: Senders' beliefs \tilde{b} on responders' back-transfers.

Received $3a$	3	6	9	12	15	18	21	24	27	30	AV
Germany											
Back-transfers	1.13	2.33	3.47	4.87	5.93	7.47	8.67	10.07	11.20	12.53	6.77
Beliefs	0.93	1.87	2.93	4.27	5.60	6.87	7.93	9.27	10.67	11.80	6.21
Israel											
Back-transfers	0.93	2.13	3.40	4.93	5.93	7.47	8.93	10.47	11.93	13.33	6.95
Beliefs	1.20	2.53	3.80	4.80	6.20	7.47	8.60	9.60	11.27	10.00	6.55
Palestine											
Back-transfers	1.60	3.33	5.00	6.40	7.47	7.67	9.47	10.73	12.00	13.87	7.75
Beliefs	1.93	4.27	6.47	8.53	11.53	13.27	16.40	18.47	21.60	23.33	12.58

Table A3: Responders' back-transfers b and senders' beliefs \tilde{b} on responders' back-transfers to fellow countrymen.

Received $3a$	3	6	9	12	15	18	21	24	27	30	AV
From Germany											
To Germany	0.93	1.87	2.93	4.27	5.60	6.87	7.93	9.27	10.67	11.80	6.21
To Israel	1.00	1.93	3.27	4.20	5.33	6.60	7.93	9.07	10.47	11.53	6.13
To Palestine	1.07	2.13	3.67	4.80	6.27	7.40	8.67	9.73	11.00	12.40	6.71
From Israel											
To Germany	1.40	2.53	3.87	5.13	6.60	7.73	7.93	8.80	10.00	10.93	6.49
To Israel	1.20	2.53	3.80	4.80	6.20	7.47	8.60	9.60	11.27	10.00	6.55
To Palestine	0.73	1.73	2.73	3.40	4.33	5.13	6.00	6.07	7.80	8.33	4.63
From Palestine											
To Germany	2.00	4.40	6.80	9.47	11.07	13.07	15.67	19.40	21.40	24.07	12.73
To Israel	1.73	3.67	5.67	7.53	9.33	11.53	15.33	17.20	20.07	21.73	11.38
To Palestine	1.93	4.27	6.47	8.53	11.53	13.27	16.40	18.47	21.60	23.33	12.58

Table A4: Senders' beliefs \tilde{b} on responders' back-transfers b .

Received $3a$	3	6	9	12	15	18	21	24	27	30	AV
Germany											
Back-transfer G-G	1.13	2.33	3.47	4.87	5.93	7.47	8.67	10.07	11.20	12.53	6.77
Belief G-I	1.07	2.13	3.67	4.80	6.27	7.40	8.67	9.73	11.00	12.40	6.71
Belief G-P	1.00	1.93	3.27	4.20	5.33	6.60	7.93	9.07	10.47	11.53	6.13
Israel											
Back-transfer I-I	0.93	2.13	3.40	4.93	5.93	7.47	8.93	10.47	11.93	13.33	6.95
Belief I-G	1.40	2.53	3.87	5.13	6.60	7.73	7.93	8.80	10.00	10.93	6.49
Belief I-P	0.73	1.73	2.73	3.40	4.33	5.13	6.00	6.07	7.80	8.33	4.63
Palestine											
Back-transfer P-P	1.60	3.33	5.00	6.40	7.47	7.67	9.47	10.73	12.00	13.87	7.75
Belief P-G	2.00	4.40	6.80	9.47	11.07	13.07	15.67	19.40	21.40	24.07	12.73
Belief P-I	1.73	3.67	5.67	7.53	9.33	11.53	15.33	17.20	20.07	21.73	11.38

Table A5: Within-country back-transfers b and beliefs \tilde{b} on foreign-country back-transfers.

Received $3a$	3	6	9	12	15	18	21	24	27	30	AV
From Germany											
To Germany	1.13	2.33	3.47	4.87	5.93	7.47	8.67	10.07	11.20	12.53	6.77
To Israel	1.07	2.33	3.40	4.67	5.73	7.00	8.20	9.47	10.60	12.07	6.45
To Palestine	1.13	2.40	3.67	4.93	6.20	7.53	8.80	10.27	11.33	12.60	6.89
From Israel											
To Germany	0.93	2.07	3.07	4.33	5.60	6.87	8.00	9.67	10.87	12.00	6.34
To Israel	1.06	2.23	3.68	5.16	6.32	7.74	9.42	10.87	12.56	13.81	6.95
To Palestine	0.87	1.87	2.87	4.33	5.47	6.67	8.00	9.27	10.53	11.67	6.15
From Palestine											
To Germany	1.73	3.60	5.13	7.53	8.67	9.33	10.93	13.67	15.33	16.93	9.29
To Israel	1.73	3.20	4.33	6.93	7.73	9.13	10.53	12.07	12.13	13.80	8.16
To Palestine	1.60	3.33	5.00	6.40	7.47	7.67	9.47	10.73	12.00	13.87	7.75

Table A6: Responders' average back-transfers b across countries.

Appendix B: Figures

Back-transfers

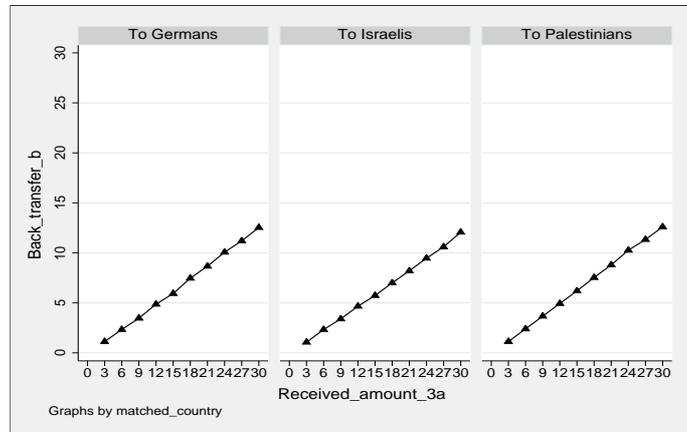


Figure A1: German back-transfers b .

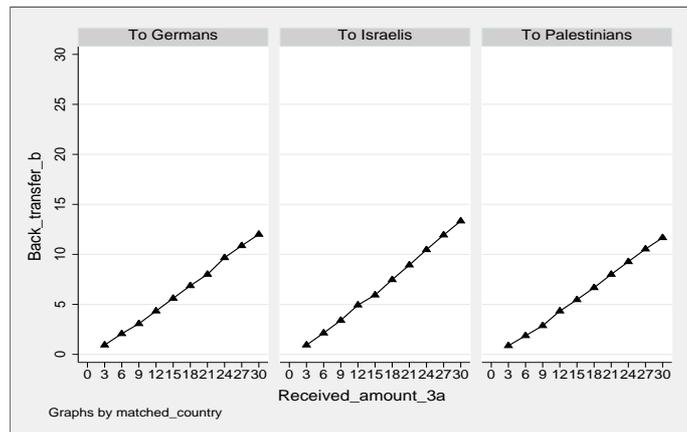


Figure A2: Israeli back-transfers b .

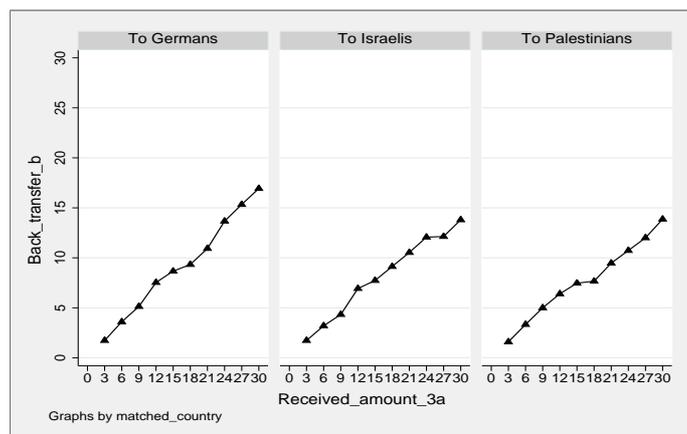


Figure A3: Palestinian back-transfers b .

Beliefs on back-transfers

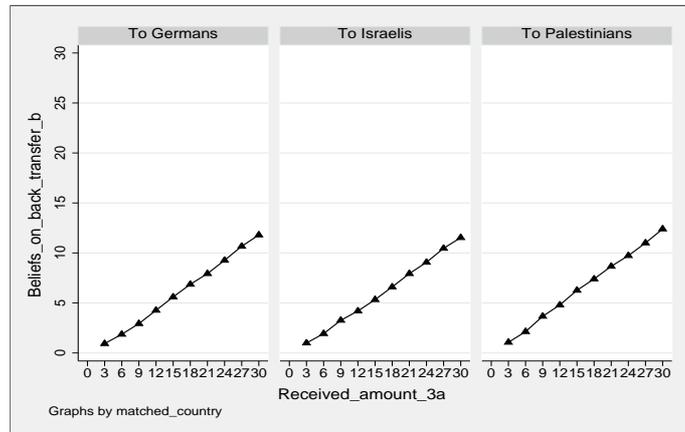


Figure A4: German beliefs \tilde{b} on back-transfers.

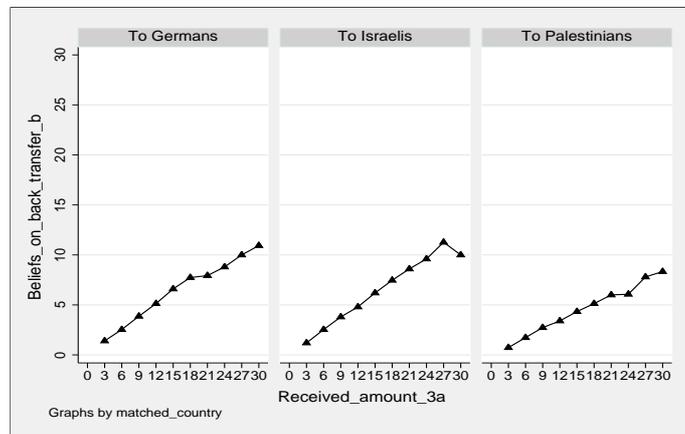


Figure A5: Israeli beliefs \tilde{b} on back-transfers.

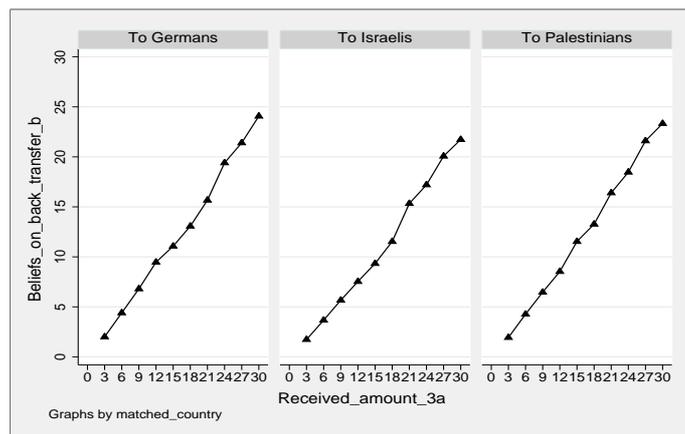


Figure A6: Palestinian beliefs \tilde{b} on back-transfers.

Appendix C: Instructions

1. Introduction

Thank you for participating in this international experiment. Please read the following instructions very carefully. If you have any questions, please raise your hand. We will then answer your question in private at your desk. During the experiment we will not answer any of your questions. In case you still have questions, please read these instructions again. During the whole experiment it is very important that you do not communicate with any of the other participants. We therefore ask you not to talk to each other.

To assure your anonymity and confidentiality you will have to randomly draw a code before the experiment starts. When presenting this code you will later on receive your payments from this experiment. This procedure ensures that we are not able to assign answers and decisions made in the experiment to specific persons.

In this experiment you can earn money. Your individual earnings depend on your own decisions and on the decisions of other participants. You will get your personal payoffs on December XX to XX, from XX o'clock to XXo'clock in room XXXX. Please watch over your code name and your code number carefully. You will need them in order to collect your payment. Without code name and code number we cannot pay out your earnings.

Right after the experiment, you will receive a fixed fee of X US\$ independent of the decisions you take in today's experiment. This fixed fee is paid in addition to the money you earned in the experiment.

During the experiment, you are involved into an interaction with a randomly assigned other participant. This randomly assigned participant takes his decisions at a different point in time than you do without having knowledge about your decisions. You are not told who this other participant is, neither during the experiment nor at any point in time after the experiment. The persons who participate with you in the experiment today are not assigned to you.

After this introduction, we will ask you some questions to check your understanding of the experimental rules. Then the experiment itself will start. After the experiment, we will ask you to fill in a questionnaire we need for a statistical analysis of the experimental data.

2. The general design of the experiment

During the experiment, pairs of two participants are formed who interact. These participants are called person A and person B from now on. All persons in one room are either persons A or persons B. Thus, persons A and persons B cannot be in one room at the same time. All persons who participate with you in the experiment at this moment belong to the same type of person.

The experiment consists of three decision rounds structured identically. Within these rounds, each participant takes decisions keeping the same role, either as person A or as person B.

At the beginning of each round, each person A and each person B receives an initial endowment X of 10 ECU (Experimental Currency Unit), the exchange rate being $1 \text{ ECU} = 0.50 \text{ US\$}$. Person B can keep this money. This amount will be paid to her at the end of the experiment, together with the payment of her experimental earnings.

Person A can now choose to either send an integer part, all or nothing of his initial endowment X to person B. This means, he can choose any integer a from the interval $[0,10]$ of his initial endowment he wants to send to person B. Each amount a chosen by person A is tripled. Thus, person B does not receive the amount a originally sent by person A but she receives three times as much, i.e. $3a$.

Here is an example: If person A sends 4 ECU, person B receives 12 ECU. If person A sends 6 ECU, person B receives 18 ECU.

During the experiment, Person B is not told which amount a person A actually chose. Therefore, person B does not know how large $3a$ actually is. Person B has to decide how much she wants to transfer back to person A. Since person B does not know the actual amount a or $3a$ respectively having been sent by person A she has to decide for each possible value of the tripled amount $3a$ how much she wants to transfer back to person A. The amount person B transfers back to person A is b . Person B can decide whether she wants to transfer back an integer part, all or nothing of $3a$ to person A. The amount b transferred back is not multiplied. Person A receives no information about person B's actual decisions during the experiment.

Person B's decision table for one round looks as follows:

Person A sent you a :	You receive $3a$:	Which amount b do you want to transfer back?
0	0	
1	3	
2	6	
3	9	
4	12	
5	15	
6	18	
7	21	
8	24	
9	27	
10	30	

After B's decision on b , the first decision round has finished for both persons. The general procedure of the following second and third decision rounds corresponds to the procedure of the first decision round.

Actual payments of the experimental earnings of person A and person B are made after all experimental sessions have been finished. By assigning person A's actual decision a to person B's decision b being valid for person A's actual decision a , individual payment for each round can be calculated.

Person A's individual profit G_A for one round is:

Initial endowment X
- amount a sent
+ amount b transferred back by person B
= G_A, Person A's individual profit for one round

Person B's individual profit G_B for one round is:

Initial endowment X
+ tripled amount a , sent by person A
- amount b transferred back by person B to person A
= G_B, Person B's individual profit for one round

Here is an example: Person A decides to send 4 ECU to person B in one round. Then, Person B receives 12 ECU. Person B decides to transfer 5 ECU back.

Person A's individual profit G_A for this round is:

Initial endowment X	10 ECU
- amount a sent	-4 ECU
+ amount b transferred back by person B	+5 ECU
= G_A, Person A's individual profit for this round	=11 ECU

Person B's individual profit G_B for this round is:

Initial endowment X	10 ECU
+ tripled amount a , sent by person A	+12 ECU
- amount b transferred back by person B to person A	-5 ECU
= G_B, Person B's individual profit for this round	=17 ECU

The individual total experimental payoff for each participant is the sum of payoffs in all three rounds plus the show up fee of X US\$ you will be paid today.

3. The experimental procedure

In the following, we precisely explain how the experiment proceeds and which consecutive decisions you will have to take.

Please read this instruction very carefully. We then will answer your questions in private. Please raise your hand if you have any questions. Before the experiment starts, we again will ask all participants whether they fully understood the instructions of the experiment.

We then will give you a short exercise to check your understanding of the experimental rules. After having solved these exercises, the experiment will start.

Start of the experiment

At the beginning, each participant randomly draws a code name. You will need the code name for being paid your experimental earnings. Furthermore, it guarantees full anonymity. Please watch over this code name very carefully.

Then, each participant randomly draws a decision sheet for each decision round. Each decision sheet contains your personal code and information about persons A or B assigned to you. Please note that the assignments as well as the order of assignments of the persons you are interacting with are randomized. On the slip containing your code you find an empty space. Please fill in your code name there.

Decision procedure for person A:*Round 1*

You learn that you are person A.

Person A is paid an initial endowment of $X=10$ ECU by the experimenters.

Person A learns that a person B was randomly assigned to him by the experimenters. Person A learns the name of the university, the city, and the country where the assigned person B is studying. With this participant, person A is interacting during this round.

Person A decides which amount a he actually wants to sent to the randomly assigned participant B. He can chose any integer from the interval $[0,10]$.

After having decided on a , the first decision round for person A has finished.

Round 2

The procedure of this decision round for person A is identical to that of round 1. However, in this round a different interacting person B is randomly assigned to person A. This means that in this second round person A does not interact with the same person B from round 1.

Round 3

The procedure of decision round 3 for person A is identical to the procedure of the previous rounds 1 and 2. However, again a different interacting person is randomly assigned to person A. This means that in this third round, person A does not interact with one of the persons B in rounds 1 or 2. During the whole experiment, each person interacts with an assigned person B only once.

The experiment has now finished for person A.

Decision procedure for person B:*Round 1*

You learn that you are person B.

Person B is paid an initial endowment of $X=10$ ECU by the experimenters. Person B owns this amount and keeps it. She will be paid this amount later, together with her experimental earnings.

Person B is told that a person A was randomly assigned to her by the experimenters. Person B learns the name of the university, the city, and the country where the assigned person A is studying. With this participant, person B is interacting during this round.

Person B is asked which amount she wants to transfer back to person A for each possible value of $3a$. Person B indicates an integer for all 11 possible amounts of $3a$ being the amount b person B wants to transfer back to person A.

After having decided on b , the first decision round for person B has finished.

Round 2

The procedure of this decision round for person B is identical to the procedure of round 1. However, in this round a different interacting person A is randomly assigned to person B. This means that in this second round person B does not interact with the same person A from round 1.

Round 3

The procedure of decision round 3 for person B is identical to the procedure of the previous rounds 1 and 2. However, again a different interacting person is randomly assigned to person B. This means that in this third round, person B does not interact with one of the persons A in rounds 1 or 2. During the whole experiment, each person interacts with an assigned person A only once.

The experiment has now finished for person B.

At the end of the experiment we would like you to answer some questions we need for the statistical analysis of this experiment.

Appendix D: Experimenting Over a Long Distance - A Method to Facilitate Inter-Cultural Experiments*

Introduction

When planning interactive experiments with geographically or administratively separated subject pools researchers face large logistic, financial, and technical problems like incompatibility of computer equipment and software (if computers are available at all), high traveling expenses and communication costs, impeded or lacking accessibility of certain subject pools. As a consequence, subject pool selection might be biased toward populations where experimentation facilities are readily available. Thus, subjects that are not easily accessible might be excluded from experimental studies.

Globalization, however, also induces interaction with populations where western technical standards are not met. To account for this fact in experimental investigations, a method is needed allowing to incorporate subject pools that are technically difficult to access. This method should enable easy-to-run interactive experiments in a true international and inter-cultural setting ¹ leading to better mutual understanding and advice for practical work.

It is known from the literature that people hold different country-specific or culture-specific dispositions. Along with an individual's perception of other cultures these dispositions influence their behavior in inter-cultural decisions (HOFSTEDE, 2001). Most experimental studies comparing behavior in different cultures are cross-cultural investigations ². Almost no studies involve true inter-cultural experiments (see however BOARINI, LASLIER, and ROBIN, 2006; BORNHORST, ICHINO, SCHLAG, and WINTER, 2007 ³).

To overcome the deficiencies mentioned above we developed a new method for Experimenting over a Long Distance (ELD in the following) that allows to simultaneously run decentralized interactive experiments in geographically separated subject pools by standardized procedural

*This part is based on the paper: "Experimenting Over a Long Distance - A Method to Facilitate Inter-Cultural Experiments" by GARI WALKOWITZ, CLEMENS OBERHAMMER, and HEIKE HENNIG-SCHMIDT (2004).

¹In the following we define experiments as intra-cultural if behavior within one culture is examined. Cross-cultural experiments are those where behavior in two or more cultures is compared with no direct interaction of subjects belonging to different cultures. Inter-cultural experiments investigate behavior of subjects in different cultures interacting directly with each other.

²See for instance WILLINGER, KESER, LOHMANN, and USUNIER, 2003; HENNIG-SCHMIDT, LI, and YANG, 2008; ANDERSON, RODGERS, and RODRIGUEZ, 2000; HENRICH, BOYD, BOWLES, CAMERER, FEHR, GINTIS, and McELREATH, 2001; ROTH, PRASNIKAR, OKUNO-FUJIWARA, and ZAMIR, 1991; see also the survey by OOSTERBEEK, SLOOF, and VAN DE KUILEN, 2004.

³The subject pool of this experiment consists of Ph.D. students all studying at the European University Institute (EUI) in Florence.

protocols. ELD minimizes logistic effort and on-site technical requirements as well as travel expenditures. It is particularly suited for simple games.

The paper is organized as follows. The next section describes the features of the ELD-method in detail. Section 3 discusses the method and concludes.

The ELD - Method

In separated subject pools, non-interactive one-shot single-player experiments as well as one-stage multi-player experiments can easily be handled because subjects in fact do not interact. Organizational problems arise, however, in multi-person multi-stage designs at different locations where participants have to take sequential decisions under time restrictions imposed by time zones and disposable resources. The crucial constraint in these experiments is the interdependence of consecutive decisions: One subject's choice depends on a prior decision of another person. ELD is designed to overcome this restriction by removing decision interdependence without creating incentive biases.

The usual sequential protocol for running investment game experiments is as follows. First, a sender and a responder get an endowment X . The sender can invest any integer $a \in \{0, 1, \dots, X - 1, X\}$ that is transferred to an anonymously matched responder. Each possible amount a chosen by the sender is tripled by the experimenter, the responder receiving $3a$ for each a . Then the responder freely decides to transfer any integer $b \in [0, 1, \dots, 3a - 1, 3a]$ back to the sender. Being dependent on the sender's action, the responder cannot make her decision without receiving information on the sender's choice. We solve this problem of sequential interdependence by applying the following features:

1. *Strategy method* (SELTEN, 1967): This method allows to organizationally disconnect the second stage of the game from the first stage. By having the responder indicate her choices for all possible sender's decisions, the sequential two-person two-stage game is converted into a two-person normal-form one-stage game for each subject. These correlated games can be played independently at different locations and different points in time. For applications of the strategy method see GÜTH, SCHMIDT, and SUTTER, (2003); BELLEMARE and KRÖGER, (2007).
2. *Pen-and-paper*: By using pen and paper, the experimental design becomes independent of equipment and software compatibility, thus reducing start-up costs. Moreover, experiments can be run in non-lab environments like classrooms. Pen and paper creates transparency with regard to procedures enhancing the credibility for participants not used to experimental methods.

3. *Remote-control organization*: When applying ELD, a central unit, the Chief Experimenter (CE) is responsible for overall planning and controlling the inter-cultural experiment. Local experimenters (LEs) are in charge of organizing and running the sessions at the locations involved (see Figure A7)⁴. CE briefs LEs in advance by a detailed procedural script and an extensive instruction manual to ensure equivalent experimental conditions in all locations involved. Instructions and decision sheets, identical up to translation into the respective languages, are prepared by CE and shipped to the corresponding LEs. Before running the experiment, CE needs special information on each participant necessary to ex-ante code and prepare the decision sheets and to randomly match players across subject pools.

At each location, participants randomly draw an ex-ante coded decision sheet before the experiment starts. Players learn about their counterpart’s pool affiliation. According to research interests, information on additional characteristics and variables is provided. All sessions in all locations having been finished, CE collects all relevant information, computes all payoffs and transfers this information to all LEs. Figure 1 gives a schematic outline of the steps required when ELD is applied to three subject pools. Possible intra-cultural assignments are not shown (see point (ii) below).

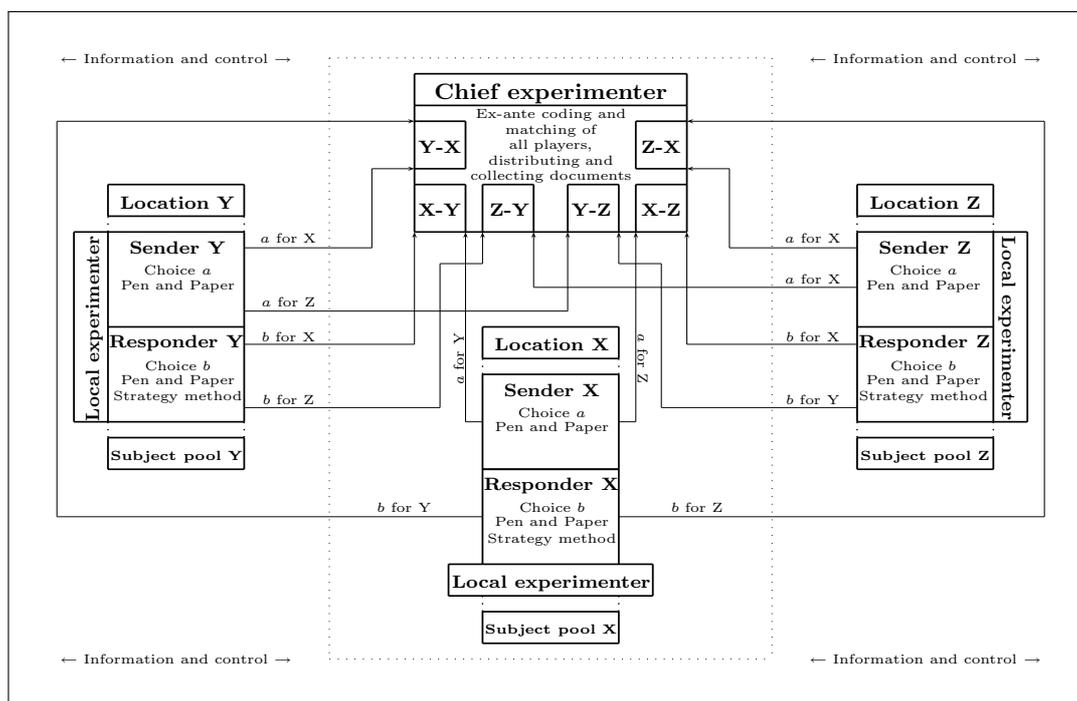


Figure A7: ELD outline for the investment game with three subject pools.

⁴In addition to saving travel expenditure having the experiment run with local experimenters avoids self presentation effects (face saving) which are likely to occur with inexperienced subjects, especially in Asia (BOND and HWANG, 1986).

Besides organizational requirements other methodological aspects are important:

- (i) To ensure equivalence of instructions, they are to be translated into the languages of the participating locations using the back-translation method (BRISLIN, 1970). If possible, CE should run pilot experiments with subjects of countries involved in the subsequent experiments. This can help to remove inconsistencies and misunderstandings of instructions and procedural features of the experiment.
- (ii) Inter-cultural comparison requires an intra-cultural control with at least two intra-cultural sub-pools to enable distinction between intra- and inter-cultural behavioral variations. Subjects may not only behave differently toward counterparts from various countries/cultures. They may also differentiate between members of their own and other subject pools within their home country/culture (e.g., FERSHTMAN and GNEEZY, 2001).
- (iii) Each subject's minimal number of decisions is equal to the number of participating subject pools when intra-personal comparison is required. With different player types and each subject having to decide in the role of *all* player types, the number of each player's decisions extends to the number of subject pools times the number of player types.
- (iv) For each choice, a separately accessible decision sheet has to be provided in order to prevent subjects from changing prior decisions or retrieving information from previous choices.

Discussion and conclusion

ELD proved to be a powerful method for running interactive inter-cultural experiments. By applying standard experimental methods in a new way ELD allows to combine intra-cultural and inter-cultural research and enables actual inter-cultural interaction. ELD can be applied to any number of subject pools and participants. Restrictions as to computer-equipped laboratories are avoided as are costs for programming. Since sessions are decentralized they do not need to be run simultaneously thus circumventing time constraints. Initiation costs are low because decisions sheets are shipped by mail.

ELD's potential drawbacks should not be neglected either. ELD can best be applied to simple experiments with few decision stages. When a large number of subject pools is involved

the resulting high number of choices might cause spill-over effects in a way that early decisions influence later choices. The pen-and-paper method does not permit computerized data collection and thus is prone to calculation and matching errors.

The strategy method may induce different behavior compared to subjects in a 'hot' environment (see, however, BRANDTS and CHARNES, 2000). Furthermore, the decision protocol may be rather complex to subjects not used to participating in experiments. On the other hand, having subjects think about every possible choice and the implications thereof requires intensive reasoning on their part. We, therefore, believe the choices to be valid and reliable data.

Given our research agenda, ELD proved an important research method in particular when the usual experimental protocol cannot be applied or turns out to be highly costly. Applying ELD we are able to conduct real inter-cultural experiments that would have been extremely difficult - if not impossible - to run as a computer experiment.

Chapter IV

Trust and Discrimination in the Labor Market - An Experimental Study with Criminal Offenders*

1 Introduction

The handling of previously convicted job applicants and employees in the labor market is a socially relevant but very complex and hardly tangible problem. At present, about 12 million ex-felons are at least potentially available in the U.S. labor market which represent 4 to 8% of civilian labor force (UGGEN and MANZA, 2002); approximately 600,000 prison inmates are released per year (LYNCH and SABOL, 2001). Hence, previously convicted represent an important social group that can dramatically alter the conditions of supply and demand in local labor markets but that is, by contrast, often overlooked (WESTERN, KLING, and WEIMAN, 2001; ROSE and CLEAR, 1998; SABOL and LYNCH, 1998). It's security of livelihood and prospect of economic improvement has sociological as well as economic significance - not only for the affected subjects themselves but for the whole society. To reduce enormous social costs caused by an inferior contribution to the level of activity in the whole economy (e.g. by unrealized taxes because of lost earnings) it appears to be relevant that feasible and effective social and economic policies to tackle job market prospects of previously convicted will be pursued and undertaken. However, the effectiveness of such policies depends on our explicit knowledge about the determinants of ex-offenders' inferior labor market performance.

While the aggregate effects of conviction on economic inequality are possibly very large, causal inference about the effect of imprisonment on labor market performance is a difficult scientific problem. An important question is whether the presented strong negative correlations reflect causation; that is, whether being convicted actually causes the former arrestee's employment opportunities and earnings to fall. And, what kind of underlying causes can be diagnosed for this evidence of disadvantages? Although there exists a vast body of literature on the connection between crime, employment, and income causal inference about the negative effect of criminal activity remains blurred. In this chapter we experimentally investigate whether

*This chapter is based on the paper: "Trust and Discrimination in the Labor Market - An Experimental Study with Criminal Offenders" by ARMIN FALK and GARI WALKOWITZ, *Working paper*, University of Bonn.

the unexplained shortfall in previously convicted people's employment probability and income level - after stripping out all observable differences between workers - exists because they are less trusted and expected to be less trustworthy¹. Conducting a controlled trust experiment in Germany with job seeking criminal offenders and not previously convicted players both acting in the role of job applicants/employees - giving their application profiles to the employers and, when employed, receiving a single wage and exerting an individual effort level - and students deciding as employers - taking a hiring and wage payment (trusting) decision we are able to generate an appropriate dataset by systematically varying our treatment variables. Our experimental design not only allows us to answer a set of relevant counterfactual questions as - What employment opportunities has a previously convicted female applicant if she would not be female, or, alternatively, not be previously convicted, or both? - but also to compare employers' attitudes towards ex-offenders with their behavior against other potential target groups of discrimination as foreigners (in this study: Turks, who represent the biggest group of immigrants in Germany²), women, or elder workers and directly contrast the degree and the quality of discrimination effects found³. The comparison of discriminatory behavior conditioned on potential targets of discrimination delivers useful insights on the interplay of factors owned by groups exposed to discrimination.

Our results give substantial support for an arbitrary disadvantage of ex-offenders. They do not only have lower chances for employment but they are also discriminated through getting significantly lower wage payments. Moreover, we find associated beliefs about employees' reciprocity to not play a primary role in interactions with ex-offenders. Consequently, as a first main result we are able to show that the reluctance to hire ex-offenders only partly consists of an inferior belief concerning their trustworthiness. Contrary, we find that it can be mainly explained by a taste for discrimination, or mere avoidance of interaction. When comparing employers' behavior toward different potential target groups of discrimination we further detect that the reluctance to hire ex-offenders and to pay them a fair wage is considerably stronger than employers' aversion to hire Turkish - a group that also faces noticeable discrimination in hiring and income - male, or elder applicants. A third important finding of this study is that wage discrimination does not occur due to a general norm. In

¹In the following we use the terms 'previously convicted', 'prison inmate', 'ex-offenders', etc. as synonyms.

²At the beginning of the 1970s, there was a particular push for the recruitment of Turkish workers in Germany. Their number rose from 172,400 in 1967 to 910,525 in 1973. The Turkish population - also as a result of family reunification - now forms the largest foreign minority in Germany. Turkish workers perform primarily "blue-collar" jobs with low payment and many of these entail high physical workloads (see GOLDBERG, MOURINMO, and KULKE, 1996).

³See for related works on race discrimination BERTRAND and MULLAINATHAN (2004), on gender discrimination WEICHSELBAUMER (2004), race and gender discrimination DARITY and MASON (1998), and on age discrimination ROSCIGNO, MONG, BYRON, and TESTER (2007).

fact, it is conditioned by employers' individual wage payment dispositions - across all target groups. This indicates that not all ex-offenders' necessarily receive low wages and that their wages are not always lower compared to wages paid to not previously convicted workers. Conducting a final individual analysis on the nature of discrimination to explore differences in discriminatory behavior, we can show that employers discriminate differently. We find discrimination against ex-offenders to be basically taste based. Contrary, the disadvantage of Turkish or elder workers mainly consists of avoidance.

The remainder of this chapter is organized as follows. In the next section two, we give an overview on the existing literature on the labor market performance of ex-offenders. Section three introduces the challenges and features of our experimental study and formulates a working hypothesis. Section four displays our results. In the last section five we discuss our results and conclude.

2 Ex-offenders and the labor market - related literature

Despite intuitive plausibility, the connection between crime and employment perspectives as well as income opportunities remains unclear. However, the negative effect of criminal conviction on employment and, primarily, on earnings has been confirmed by many studies and across data gathering methods. Analyzing survey data WESTERN (2004) shows that men who have been incarcerated have significantly lower employment rates and annual earnings compared to men who have never been incarcerated. HOLZER, RAPHAEL, and STOLL (2002) found that over 60% of employers indicate that they would 'probably not' or 'definitely not' be willing to hire an applicant with a criminal record. This aversion is considerably stronger than the aversion of employers to hire applicants from other commonly stigmatized groups of workers (e.g., welfare recipients). WESTERN and BECKETT (1999) and FREEMAN (1992) demonstrate that incarceration is associated with a small but persistent decrease in weeks worked after release. Using employment data FREEMAN (1992) finds that ex-offenders face a 15 to 30% reduction in employment probability compared to a control group. LOTT (1990) detects very large negative effects of prison sentence length on employment and earnings, especially for high income individuals. Evidence from administrative records points the same way: Looking at total earnings from 1983-1991 KLING (1999) and NEEDELS (1996) show that an additional year of incarceration reduces total earnings by about 12%. Moreover, GROGGER (1995) finds employment rates to be 4% lower in the group of previously convicted. Taken together, these findings clearly indicate that chances for a secure (re-)employment with stable and appropriate income are comparably low for former criminal offenders, especially for young males who represent the majority of released prison inmates.

The existing literature delivers different explanations for the labor market effects of imprisonment: First, ex-offenders suffer from the stigma of criminal conviction. A criminal record delivers an observable signal that is usually correlated with low expectations in productivity or reliability of previously convicted applicants (PAGER, 2003; BOSHIER and JOHNSON, 1974; BUIKHUISEN and DIJKSTERHUIS, 1971; SCHWARTZ and SKOLNICK, 1962). Second, their status makes them less employable, as employers fear the legal liabilities that could potentially be created by hiring ex-offenders. Federal or state law even forbids the hiring of convicted felons or ex-offenders into some sectors, such as interstate transport, finance, and child or patient care. More broadly, employers seem to fear legal liability where ex-offenders have to deal directly with customers or handle property that belongs to others (HOLZER, RAPHAEL, and STOLL, 2001). Third, vacant jobs are often assigned through social networks where previously convicted people do not take part (GRANOVETTER, 1995; SULLIVAN, 1989). Previous ties to these networks break off while imprisoned and new contacts are frequently established between prison inmates which are more likely to lead again to criminal activity after release and seriously impede a successful reintegration into stable social structures (VENKATESH, 2000; SÁNCHEZ-JANKOWSKI, 1991). Fourth, in order to explain lower labor market opportunities for former prison inmates many authors argue that during the period of imprisonment their social capital is eroded which is reflected in a lack of actual and expected productivity. While imprisoned, possibilities to acquire new skills and expertise are limited. Human capital, as far as it is already existing, can only partly be strengthened or improved inadequately (WESTERN et al., 2001; WALDFOGEL, 1994a,b). It rather appears that the probability for acquired skills to wear off is very high. The time spend in a penitentiary also seems to aggravate mental and physical disease patterns which already appeared before imprisoned (HOLZER et al., 2002). Consequently, during job-seeking and wage determination ex-offenders usually compete against other, not previously convicted applicants who in general are not effected by the named disadvantages.

An alternative claim is made by HOLZER et al. (2002). They argue that from the viewpoint of employers, a criminal history record may signal an untrustworthy or otherwise problematic employee. Employers may avoid such workers due to a perceived increased propensity to break rules, steal, or harm customers. To make investments in skills and training, employers must trust their workers to stay with the firm and repay the investment. Likewise, workers must trust managers to provide job security and a career ladder, particularly early on if primary sector wages are lower than those offered in the open labor market. AKERLOF (1982) has described this ethic of mutual obligation as a partial gift exchange, where hard work and loyalty are exchanged for job security and high wages. Hence, trust and reciprocity display

constitutive mechanisms in every bilateral employer-employee relationship⁴.

Examining income panel data and administrative records WALDFOGEL (1994a) explored the causes of the conviction effect and how it varies with characteristics of offense and offender, in particular whether an offender was in a position of trust prior to conviction. He finds that conviction particularly stigmatizes those who used the trust, granted him or her in a responsible or official capacity, to violate the law. However, his study is partly based on pre-sentence data retrospectively prepared by probation officers and post-conviction data from monthly probation reports. Probation officers were instructed by a coding manual to indicate whether “*the offender used the trust received by virtue of his employment to commit the offense*” (WALDFOGEL, 1994a, p. 65). Thus, previous employers were not directly involved; the classification of jobs and associated required trust (in different branches) was executed by exterior probation officers. In addition, the impact of a breach of trust on employment probabilities and earnings was indirectly estimated by differences in differences. One innovation of our study is that we can control for a homogeneous composition of our employers subject pool having German students holding a similar demographical background deciding as employers. In addition, we are able to directly observe and measure employers’ trust and beliefs executing truly exogenous *ceteris paribus* changes. Moreover, using a straightforward trust-reciprocity mechanism we can endogenously control our dependent variables as described in the next section.

3 An experimental approach

In this section we introduce and describe our experimental setup. Firstly, we discuss the advantages of conducting a laboratory experiment to find an appropriate answer to our initial research questions. Secondly, we address some methodological specifications concerning causal inference and illustrate our experimental design. Then we explain our selected treatment variables. After that, the experimental procedure is presented in detail. In the last part, based on previous studies and the strategies provided in our experiment, we state a working hypothesis.

3.1 Why a laboratory experiment?

Evidence on employment and wages, gained so far from polls and labor market data, appears to be partly inconsistent and is subject to several methodological problems. In many cases,

⁴Contracts between employers and employees are often incomplete, i.e., the employer is neither able to agree *ex ante* on the employee’s actions by contract nor to observe and pay or sanction all associated actions.

important factors confounded with a treatment are not observed and recorded. Differentiated statements about the intensity of the correlation between conviction, re-employment, and income level as well as about its causes are very vulnerable.

Previously convicted people's economic activity often falls within a gray area that escapes detection by the criminal justice system and social insurance agencies. Their housing and patterns of household attachment often place them out of the reach of traditional social survey methods and thus complicate access to data scientifically utilizable. Furthermore, the risk that personnel managers' verbal comments and decision data collected in previous studies are biased because of selection, framing, or social desirability effects exists. The questionnaire approach used in many earlier studies seemed biased to the problem since respondents confronted with hypothetical situations might be particularly prone to answer in what they considered a socially acceptable manner. Conducting an experiment we can bring employers and workers to the laboratory and let them take incentivised decisions.

Another common issue conductors of field experiments on discrimination (e.g., audit studies) are confronted with is unknown heterogeneity of the labor supply side. Among others, employers might vary in sex, age, race, branch, income, personality, values, actual mood, and in variables unobservable to the researcher that might guide their behavior in the very moment of decision. Consequently, it is very difficult to involve a homogeneous and comparable group of employers. This issue is often disregarded. Bringing students to the laboratory enables us to investigate a very homogeneous subject pool in the role of employers. Typically, students have low and comparable opportunity costs. Moreover, although there might be some quantitative differences compared to actual employers qualitative effects are expected to be similar⁵.

In general, experiments have become an essential instrument to circumvent issues arising in field experiments and questionnaire studies and to find behavioral regularities under controlled conditions to explain labor market phenomena. Going into the laboratory permits to strictly monitor exogenous variables as well as to control endogenous parameters. As long as the relevant underlying conditions remain substantially unchanged experimental data allow causal inferences (see FALK and FEHR, 2003).

⁵For a review of this topic see LIYANARACHCHI (2007). Moreover, comparing managers' and students' trusting behavior FEHR and LIST (2004) found no fundamental differences among both groups. COOPER, KAGEL, LO, and GU (1999) showed that in a repeated game behavioral differences between managers and students decrease over time.

3.2 Methodological challenges and causal inference

Being convicted is not a randomly assigned treatment. Typically, economists model crime as an economic choice by rational agents. It is not determined by bad attitudes or mental illness, but it is made on the basis of a maximization problem in which agents have to compare costs and benefits of legal and illegal activities taking into account the probability of being arrested and punished and the expected returns from crime (BECKER, 1957). Unlike sex, race, or age which are (mostly) randomly assigned by nature to all social classes, previously convicted persons represent a highly self-selected group. To become a criminal convict requires - may it be voluntarily or not - a prior act of criminal activity. Hence, conviction to some extent delivers a signal of actual trustworthiness or productivity.

However, contributions by many authors have shown that becoming involved in criminal activity is highly correlated with socio economic factors like family or neighborhood background (see for an overview BUONANNO, 2003). Factors as civil status (unmarried) or low age of the mother, mother's teenage and unwanted pregnancy, low mother's education, low household income, low received level of education, and neighborhood with high crime rates (peer effects) can influence the probability for an individual to get in conflict with law. Furthermore, crime is partly a function of policy choice which exogenously determines detection probability and degree of punishment. In general, through the interplay of these factors disadvantaged young and unskilled men, living in suburban or high-poverty or high-crime areas, face the highest risk to fall into crime.

Consequently, any causal statement about the effects of incarceration must consider the rival explanation that criminal offenders are highly self-selected and have few employment prospects regardless of incarceration (CASPI, WRIGHT, MOFFITT, and SILVA, 1998; MOFFITT, 1993; SULLIVAN, 1989). Why should ex-offenders get employment when they chose to get in conflict with law? Is it not rational to not select them because their pre-market skills are likely to be inferior? To disentangle this causal inference issue we have to answer the counterfactual question: What employment (income) opportunities has a previously convicted applicant if he or she would not hold a criminal record? We address this issues as follows: First, we strictly limit the field of interaction to a bilateral trust-reciprocity exchange. Second, we keep certain characteristics constant that are negatively correlated with criminal activity - namely 'civil status', 'education', 'occupational qualifications', and 'additional professional skills'. Only treatment variables' parameter values are systematically varied. Moreover, no additional information about other socio economic factors or physical/psychological conditions of applicants are delivered to the subjects.

The assignment of incarceration resembles a randomized experiment, in which the treatment group (previously convicted persons) is identical to the control group (not previously convicted persons) in all respects, except incarceration status. If applicants with conviction records differed only randomly from everyone else, we are able to infer from our data whether ex-offenders' employment chances and income levels are inferior because of the lack of trust.

Our framework and the fact that we included real prison inmates on the one hand allows us to put subjects as much as possible into a real life context. On the other hand, planning our experiment we faced the issue that previously convicted subjects were not easily accessible. Due to administrative and legal barriers and because of our specific requirements to their individual profiles the number of potential participants was very small. Consequently, we decided to re-match previously convicted applicants with different employers. In order to keep conditions constant we applied the same procedure to the control group.

3.3 Experimental design

The basic technical mechanism for our experiment is provided by the Gift-Exchange-Game as introduced by FEHR, KIRCHSTEIGER, and RIEDL (1998). At the first stage a player (a student) in the role of an employer chooses an arbitrary integer wage w , with $0 \leq w \leq 100$, which he wants to pay an assigned employee (either a prison inmate or a control group member) in advance for her working effort⁶. Subsequently, at the second stage, the player in the role of the employee, knowing the specified wage, decides how much effort she wants to contribute for that earned wage. Her arbitrary integer work effort e , with $1 \leq e \leq 10$, reduces her income, because the effort induces the costs $c(e)$. Table 1 displays the cost function of the employee.

Effort	1	2	3	4	5	6	7	8	9	10
Cost(Effort)	0	1	2	4	6	8	10	12	15	18

Table 1: Employee's cost function dependent on exerted effort.

Formally, the employer's payoff function is given by $\pi_{employer} = 10 \cdot e - w$. Thus, the employer's income depends on the wage he pays the employee and the work effort the employee exerts. The more effort the employee puts into work, the higher is the employer's income. Simultaneously, the lower the wage he pays the employee, the higher his income. Accordingly, employee's income can be stated as $\pi_{employee} = w - c(e)$. It is the wage she receives from her employer minus the costs for the work effort she has to bear. The higher the employee's

⁶In the following, we denote the employer as male and the employee as female.

wage, the higher is her income. Simultaneously, her income is higher the lower her chosen work effort is.

The in advance wage payment by the employer can be interpreted as his level of trust toward his employee, since she can freely determine her level of work effort afterwards. The chosen work effort reflects her degree of reciprocity or reliability toward her employer. With perfect information, the standard subgame perfect equilibrium in this game yields $e^* = w^* = 0$. A rational employer will pay no positive wage, since he knows that a purely self-interested employee has no incentive to exert any effort. However, if both players coordinate their actions both parties can reach higher income levels compared to the game theoretical outcome.

3.4 Treatment variables

Treatments were implemented as follows: Employers received 16 short curricula vitae (profiles) containing real information about applicants (see Figure 1).

Age	<i>24 <u>or</u> 31</i>
Sex	<i>Male <u>or</u> female</i>
Citizenship	<i>German <u>or</u> Turkish</i>
Civil status	<i>Single</i>
Education	<i>Secondary modern school qualification</i>
Occupational qualifications	<i>Completed apprenticeship</i>
Additional professional skills	<i>PC-skills, industrial arts, Basic skills in English, driving license</i>
Notes	<i>Engaged in learned profession <u>and</u> (no entry) <u>or</u> Previously convicted, conviction due to property crime <u>or</u> Previously convicted, conviction due to violent crime (no sexual offense)</i>

Figure 1: Employees' profile given to employers containing treatment characteristics.

On the left hand side of the profile characteristics are shown, on the right hand side the according parameter values are illustrated. As treatment variables we implemented 'age' (AGE), 'sex' (SEX), 'citizenship' (CIT), and 'status' (STAT, being part of 'notes'). These treatment variables were carefully chosen in advance in cooperation with representatives of the federal ministry of justice of North-Rhine Westphalia (Germany) in order to design realistic profiles and to study the interaction of the selected treatment variables. Combining the different parameter values of each treatment variable - '24' or '31' (AGE), 'male' or 'female' (SEX), 'German' or 'Turkish' (CIT), 'no entry' or 'previously convicted, conviction due to property crime' or 'previously convicted, conviction due to violent crime' (STAT) - 16 typical and realistic application profiles are emerged representing the target group of previously

convicted applicants and an otherwise identical control group⁷. The characteristics ‘civil status’, ‘education’, ‘occupational qualification’, and ‘additional professional skills’ obtained in all 16 profiles only one value and were kept constantly. We did this to give as much realistic information as possible to the subjects and to reduce possible demand effects of our treatment variables.

3.5 Experimental procedure

Meeting the particular conditions of conducting an experiment with students acting as employers and prison inmates (and the control group) being employees both stages of the game were technically and locally separated. Therefore, participants from the first and second stage acted time-delayed. In the following, the experimental protocol is described in more detail⁸.

Employers’ decisions (1st stage)

The first stage of the experiment was conducted with students of the University of Bonn. In each of six consecutive sessions 16 subjects acted in the role of employers within a generated application situation on a labor market including 16 employers and 16 applicants⁹. The computer-aided experiment was started after all participants had correctly answered a questionnaire of control questions about the experiment. Employers first received the 16 uniquely coded short curricula vitae with information about the 16 applicants who all applied at all the employers for an one period working project¹⁰. It should be stressed that employers (students) were explicitly told that applicants were real people who exactly possessed the characteristic traits as stated in their personal data sheet. Moreover, employees were selected in advance by the authors and probation officers exactly according to these profiles.

After having checked all obtained profiles subjects were asked to take three decisions: 1. Determination of an individual rank order, 2. Wage payment, 3. Belief on exerted effort. In the first step subjects were requested to set up a personal ranking order. Therefore, they had to assign each applicant a ranking position between 1 (best rank) and 16 (lowest rank) on their computer screen, with rank 1 being the applicant they preferred most for employment and rank 16 being the least preferred applicant. This ranking order was set up simultaneously by all employers on the market. By this personal ranking order the probability for an employer

⁷Turkish prison inmates did not take part in the study. Thus, out of 24 possible types of applicants, only 16 were placed in order to reduce complexity for students.

⁸For instructions please see Appendix C.

⁹At the beginning of this (laboratory) stage employees are named ‘applicants’ because they are not yet assigned to a specific employer. Furthermore, here they remain in a passive role.

¹⁰Profiles were neutrally coded by a randomized and meaningless combination of three letters. All 16 employers were equal to the applicants. Applicants received no additional information about employers. We assumed that applicants hold no specific preferences concerning *where* to apply.

to actually hire a certain applicant later on in order to interact with her was influenced. In a second step, employers decided about the integer amount of wage w , they would actually pay an applicant in case of employment. The employers again used the profiles of the 16 applicants to reach a decision within this step. Since they did not know, at this point in time, which applicant they would actually employ, they were asked to define a wage for each of the 16 applicants. Thus, every wage payment got potentially relevant. Finally, students were asked which integer amount of work effort e they expected the applicants to contribute for every wage payment.

After completion of the ranking, wage payment, and belief statements for every applicant, each employer had to draw a random number between 1 and 16. The random numbers were drawn one after another by the participants themselves without putting them back and in the order of their cabin numbers randomly determined at the beginning of the experiment. This random number was the indicator for the order of the employers' turns to actually hire an applicant from the labor market. Thus, the number also influenced the probability for an employer to be able to engage the favored applicant. The employer who drew random number 1 was coupled with the applicant he preferred most and, hence, he was allowed to actually employ the applicant he ranked best. This matched and afterwards employed applicant was from then on no longer available on the labor market for the remaining other 15 employers. Hence, she was automatically deleted from the individual ranking orders of the other 15 employers. Consequently, 15 available applicants remained on the labor market. Then it was the turn of the employer who had drawn random number 2. The applicant who was now ranked best among the remaining 15 applicants in his personal ranking order was assigned to him. She could have been the applicant the employer had originally ranked first. In case the employer with random number 1 had already engaged her, the applicant ranked second was assigned to the employer with random number 2. Afterwards it was the turn of the employer with random number 3, and so forth. This method was applied until all applicants were assigned to an employer¹¹. Thus, at the end of the session each applicant was actually assigned to an employer. This means within the first stage after every session there were 16 employer - employee pairs. After having drawn the random number and having fed it into a separate computer by the experimenters, employers were shown the corresponding code of the applicant's profile on their computer screen and hence knew which applicant they had actually employed. The wage, determined earlier by the participant, was only paid, in each case, this certain employee. Also, only the work effort of this employee (later determined in the

¹¹This matching mechanism was common knowledge at the beginning of the experiment and is incentive compatible, i.e., it is a (weakly) dominant strategy for each employer to rank applicants in accordance with his/her true preferences. Incentive compatibility is shown in Appendix A.

second stage of the experiment) was relevant for the income of the matched employer. After having completed a questionnaire, the experimental session was finished for the students. The participants received their payment immediately after the completion of the sequent second experimental stage.

Employees' decisions (2nd stage)

Within this stage, the 16 employees (former applicants) were asked to choose (in individual sessions) a work effort they were willing to offer for a given wage. The experiment was not started before a comprehensive, mainly oral, introduction to the experimental setup took place and until all participants had answered a list of control questions about the experiment correctly. First of all, in order to reach a decision, the employees were told that 6 potential employers were assigned to them¹². Additionally, every employee got to know the amount of wage, each of the 6 employers would pay her. The employees neither learned what wage any employer was willing to pay a different applicant nor whether the respective employer had preferred a different applicant more (less). The employees were then asked to choose an arbitrary integer work effort e . While employees took their decisions the experimenters left the room to guarantee anonymity and to avoid social desirability biases¹³. Finally, in order to calculate the employees' actual income, one payoff-relevant employer was determined by drawing lots out of the pool of available employers¹⁴. The employees received their payment immediately after their individual sessions. Prison inmates received as a payment a credit note on their in-house account in prison.

Location	Stage	Session	Subjects	SEX	AGE	CIT	STAT
University of Bonn	1	1-6	6x16=96	M=52/F=44	⊙ 24.4	G	Students
Individual sessions Cologne	2	7-8	1/1	M/F	24	G	NC
Individual sessions Berlin	2	9-10	1/1	M/F	31	G	NC
Individual sessions Berlin	2	11-12	1/1	M/F	24	T	NC
Individual sessions Berlin	2	13-14	1/1	M/F	31	T	NC
Penitentiary Gelsenkirchen	2	15-16	1/1	F	24/31	G	P
Penitentiary Gelsenkirchen	2	17-18	1/1	F	24/31	G	V
Penitentiary Bochum-Langendreer	2	19-20	1/1	M	24/31	G	P
Penitentiary Bochum-Langendreer	2	21-22	1/1	M	24/31	G	V

Table 2: Experimental protocol. We denote M=male, F=female, G=German, T=Turkish, NC=not previously convicted, P=property crime, V=violent crime, 24=24years old, and 31=31 years old.

During the entire experiment the monetary unit 'points' was used. At the end, all achieved points were converted into EURO with an exchange rate of 1 point = 0.10 EURO and the

¹²The number of potential employers for an employee derives from the number of experimental sessions carried out in the first stage. At the end of every session, each employer had been assigned one applicant.

¹³In addition, we explicitly guaranteed prison inmates that no decision is submitted for probation officers' attention and that we can exclude any impact on their future probation treatment.

¹⁴Employers did not know that their employees also received five other wages from different employers. However, due to organizational constraints we decided upon this procedure to maintain incentive compatibility.

amount was paid to the participants. Both employers and employees could have achieved a deficit. In case this occurred he or she had to balance it out of his endowment¹⁵. To sum up, Table 2 shows the organizational protocol of our decision experiment.

The first stage - the employers' decisions - was conducted with students in the Laboratory for Experimental Economics at the University of Bonn. 16 subjects of all different branches of study, of both sexes, and between 20 and 30 years old took part in each of the six sessions. Altogether, decision data of 96 students (52 males, 44 females) were gathered at this stage. Subsequently, four experimental sessions were conducted each in the penitentiary of Gelsenkirchen (females) and in the penitentiary of Bochum-Langendreer (men). In each individual session one prison inmate¹⁶ of different age (24 or 31 years old) and crime-profile (previously convicted because of property crime or previously convicted because of violent crime), participated. Also part of the second stage was the data ascertainment of the control group consisting of four German subjects who did not have a criminal record as well as four Turkish citizens living in Berlin or Cologne who also had an empty record. All of them had comparable demographic characteristics to the prison inmates.

3.6 Behavioral prediction

As illustrated, PAGER (2003), WALDFOGEL (1994a), BOSHIER and JOHNSON (1974), and BUIKHUISEN and DIJKSTERHUIS (1971) have argued that criminal conviction signals convicts' untrustworthiness evoking lower employment prospects and income level for ex-inmates. We conjecture an ample quantity of subjects to hold a stereotype or negative prejudice about ex-offenders' reciprocity and therefore discriminate to maximize expected payoff. Furthermore, following HOLZER et al. (2002), WALDFOGEL (1994a), BECKER (1968), and SCHWARTZ and SKOLNICK (1962), subjects might simply dislike stigmatized previously convicted applicants/employees, independently of an expectation concerning their expected effort. Since both arguments point into the same direction implying lower employment probability and limited income for previously convicted applicants/employees we can state our working hypothesis by¹⁷:

Hypothesis: *Previously convicted are paid lower wages and are less preferred for employment (assigned minor ranks). They are either expected to be less trustworthy compared to the*

¹⁵Participants received an initial endowment of 10 EURO which included the payment for the treatment of a comprehensive ex post questionnaire.

¹⁶The prison inmates who took part in our study were about to be released from prison within 3 months. They all planned to apply for jobs after release.

¹⁷We disregard the case where employers may intentionally want to hire an ex-offender to punish her later on by a low wage payment.

control group (and therefore less trusted) or simply disliked and less preferred, independent of an (existing) underlying and stated belief.

Our hypothesis is approved given that all other information about ex-offenders is kept constant in the experiment.

4 Results

In this section we will first analyze employers' hiring decisions, then chosen wages - our potential measure for employers' trust. We will then focus on employees' actual decisions and employers' expectations on exerted efforts. The last subsection explores whether we find evidence for systematic discriminatory behavior of employers toward ex-offenders, Turks, females, or older employees on the individual level. During our analysis we will denote second movers as 'applicants' when analyzing employers' hiring decisions and name them 'employees' when looking at wages. Furthermore, in order to facilitate reading we use abbreviations **PC**, **NC**, **G**, and **T** for identifying (not) previously convicted, German, and Turkish applicants/employees.

		Hiring		Wages		Expected effort		
		NC	PC	NC	PC	NC	PC	
(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	
(1)	SEX	M	5.5 (2.1)	12.2 (1.8)	31.8 (22.5)	22.1 (19.5)	4.66 (2.62)	3.23 (2.24)
(2)		F	4.8 (1.9)	11.5 (1.8)	32.7 (22.8)	22.3 (19.3)	4.90 (2.68)	3.39 (2.32)
(3)	AGE	24	4.7 (2.3)	11.7 (1.9)	32.0 (22.3)	22.0 (19.2)	4.82 (2.66)	3.29 (2.24)
(4)		31	5.6 (2.1)	12.0 (1.7)	32.5 (23.4)	22.4 (19.5)	4.74 (2.67)	3.33 (2.30)
(5)	CIT	G	4.3 (2.0)		33.1 (22.7)		4.90 (2.66)	
(6)		T	6.0 (2.1)		32.3 (22.6)		4.66 (2.64)	
(7)	STAT	P		11.2 (2.1)		22.4 (19.7)		3.33 (2.32)
(8)		V		12.5 (2.9)		22.0 (19.7)		3.29 (2.36)
(9)		Total	5.1 (1.5)	11.9 (1.5)	32.2 (22.6)	22.2 (19.3)	4.78 (2.62)	3.31 (2.25)

Table 3: Average mean ranks, wages, and beliefs (number in brackets denotes standard deviation). We denote M=male, F=female, G=German, T=Turkish, P=property crime, and V=violent crime.

Table 3 gives an overview on average ranks, wages, and expected efforts for **NC** and **PC** applicants/employees. Column (I) shows the treatment variables which are to be analyzed and column (II) the according parameter values. In columns (III) and (IV) ((V) and (VI)) average ranks (wages) for **NC** and **PC** applicants/employees are opposed to each other, for every parameter value of all treatment variables. Columns (VII) and (VIII) display employers' average beliefs on exerted efforts given the wages they indicated for a certain employee.

4.1 Hiring (rank order)

First, average ranks, which employers assigned to **NC** and **PC** applicants, are compared. **NC** applicants were assigned a significantly better rank (5.1) than **PC** applicants (11.9, $p < .01$, Wilcoxon signed rank test, two-sided, row (9), columns (III) and (IV), Table 3). This means, employers in general clearly preferred **NC** applicants within their hiring decision.

Observation H1¹⁸: *PC applicants are significantly less preferred for employment than NC.*

Figure 2 shows average ranks for all distinct profiles and the distribution of assigned ranks for **NC** and **PC**. In both figures a substantial leap from rank 8 (last rank to be assigned to a **NC** applicant when all **NC** are more preferred) to rank 9 (first rank to be assigned to a **PC** applicant) can be observed. Hence, a considerable split into **NC** and **PC** applicants (without overlap) can be determined in employers' hiring behavior.

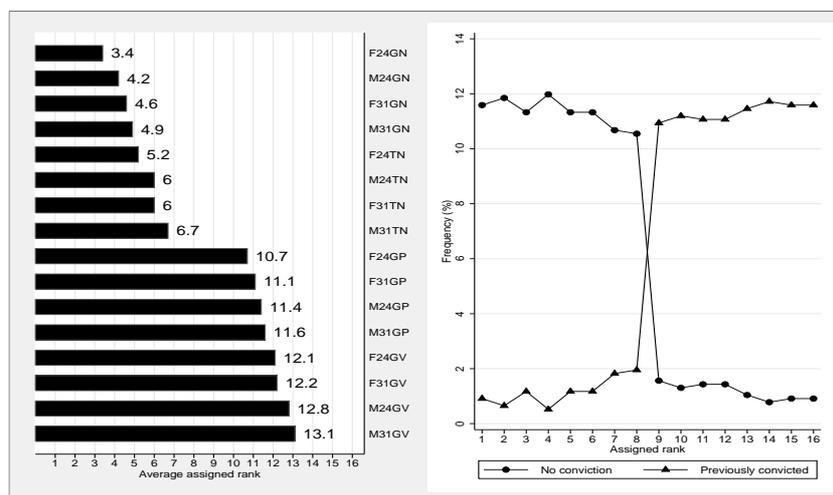


Figure 2: Average and distribution of assigned ranks.

Since a well-defined and highly significant separation of **NC** and **PC** is verifiable, both subgroups can be further analyzed in more detail separately. Among **NC** applicants (column (III), Table 3) participants significantly distinguished between **G** (4.3) and **T** (6.0) applicants ($p < .01$, rows (5) and (6)). Moreover, **females** (4.8) were clearly preferred to **males** (5.5, $p < .01$, rows (1) and (2)). Applicants who were **24 years old** (4.7) were significantly favored over elder applicants (5.6, $p < .01$, rows (3) and (4)) (all Wilcoxon signed rank test, two-sided).

¹⁸H denotes 'Hiring'.

Observation H2: *Among NC T applicants are less preferred compared to G applicants. Females (24 year olds) are more preferred than males (31 year olds).*

Within the group of **PC** applicants (column (IV), Table 3) a significant discrimination between property crime (11.2) and violent crime (12.5) can be detected ($p < .01$, rows (7) and (8)). Analogously to **NC** applicants, employers significantly preferred **female** (11.5) over **male** applicants (12.2, $p < .01$, rows (1) and (2)). However, no substantial preference for **24 year old** (11.7) or **31 year old** applicants (12.0) can be observed ($p = .277$, rows (3) and (4)) (all Wilcoxon signed rank test, two-sided).

Observation H3: *Among PC violent crime applicants are less preferred than those with a property crime record. Females are more preferred compared to males. Employers do not discriminate according to PC applicants' age.*

The detected hiring scheme gets further support by a conjoint-measurement-analysis. In addition, using this method the relationship between employers' actual choices dependent on different potential targets of discrimination (e.g., **PC** and **T**) can be examined. In a conjoint-measurement-analysis a multi-attributive preference structure model characterizes the contribution of single parameter values a (e.g., **G** or **T**) to the perceived subjective total utility of a specified profile c ¹⁹. It is assumed that employer's total utility of a profile is additively composed by the part-worth utilities of the profile's treatment variables' parameter values. In our case, a typical preference structure model for all $n = C = 96$ subjects is to be estimated. The corresponding minimization problem is given by:

$$\sum_{i=1}^{96} \sum_{c=1}^C (U_{ci} - \hat{U}_{ci})^2 = \sum_{i=1}^{96} \sum_{c=1}^C \left(U_{ci} - \left(\hat{\alpha}_{0i} + \sum_{a=1}^{A_{m-1}} \sum_{m=1}^M \hat{\alpha}_{aim} \cdot \hat{x}_{acm} \right) \right)^2 \rightarrow \min!$$

with:

\hat{U}_{ci} = estimation of total utility value that an employer i assigns to profile c

$\hat{\alpha}_{0i}$ = estimated basic utility

$\hat{\alpha}_{aim}$ = estimated part-worth utility of parameter value a from treatment variable m for employer i

$$\hat{x}_{acm} = \begin{cases} 1 & \text{if profile } c \text{ contains value } a \text{ in treatment variable } m \\ 0 & \text{otherwise} \end{cases}$$

¹⁹See for an overview: BENNA (1998); HAIR, ANDERSON, TATHAM, and BLACK (1995); BACKHAUS, ERICHSON, PLINKE, and WEIBER (1994); GREEN, TULL, and ALBAUM (1988).

Applying an OLS-estimation algorithm the obtained part-worth utilities are identical with the arithmetic mean of the corresponding part-worth utilities gained from each employer. For each parameter value combination of the treatment variables SEX, AGE, CIT, and STAT we can now determine employers' total utility for each profile²⁰. Figure 3 displays employers' estimated total utility for all 24 possible application profiles.

The outcome of our analysis ascribes the treatment variable STAT to the highest importance (63.99%) within the rank order decision with the parameter values 'not previously convicted', 'property-' and 'violence crime' being clearly distinguished. A Fisher-Pitman permutation test for paired replicates ($p < .01$, two-sided) clearly confirms estimated utility derived from **NC** profiles (11.86) to be significantly higher than from **PC** job seekers (4.28) and an explicit neglect of violent crime applicants against property delinquents (3.40 and 4.96, $p < .01$, two-sided).

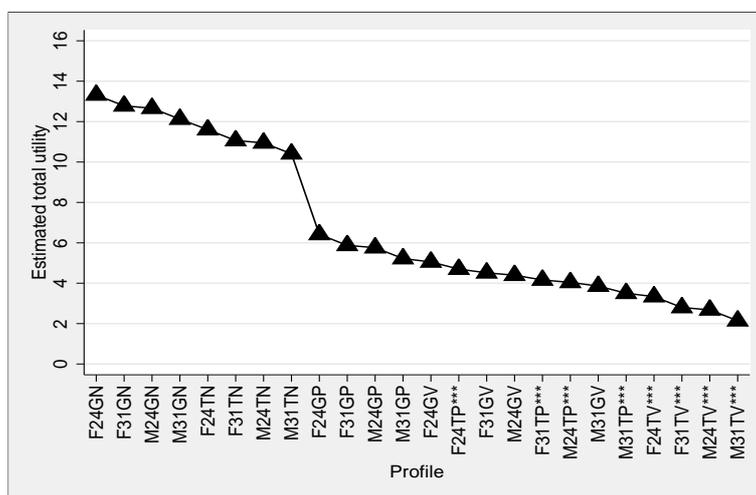


Figure 3: Estimated utility for all 24 profiles (***) denotes profiles not included in the study).

The treatment variable CIT maintains a relative importance of 13.95%, with **G** applicants being considerably preferred to **T**. Again, the test statistic shows this distinction to be substantial (7.67 and 5.95, $p < .01$, Fisher-Pitman permutation test for paired replicates, two-sided). Although, by comparison, the treatment variables AGE and SEX (relative importance of 11.06% and 10.99%) do not strongly influence employers' ranking order considering the entire group of applicants, employers in general have a higher utility from **female** compared to **male** (7.14 and 6.48) and from **24 year old** compared to **31 year old** applicants (7.08 and 6.54, both $p < .01$, Fisher-Pitman permutation test for paired replicates, two-sided). In

²⁰The total utility can also be calculated for profiles not originally part of our investigation (e.g., M31TV).

total, when we look at the relative importance of all treatment variables we find STAT to have a 4.6 (5.8) times higher impact on employers' ranking decisions than CIT (AGE and SEX). Moreover, CIT is 1.3 times more influential compared to AGE and SEX. There exists a nearly perfect correlation between estimated and empirical ranking data (Kendall's $\tau = .967$, $p < .01$)²¹.

Observation H4: *The profile treatment variable STAT has the highest impact on employers hiring decisions. The relative importance of CIT, SEX, and AGE is considerably lower.*

To conclude this subsection we summarize our findings in the following result for hiring:

Result 1: *A previous conviction negatively influences employers' hiring decisions. Most preferred for employment are 24 year old female German applicants who are not previously convicted. Contrary, most disadvantaged are those convicted older candidates with Turkish citizenship. Moreover, the STAT-signal operates significantly stronger than the treatment variables CIT, SEX, and AGE which also adequately affect employers' behavior (Turks are less and females and 24 year olds are more preferred).*

4.2 Wage payment (transfer behavior)

In the next step we compare wages employers have indicated to pay **NC** and **PC** employees. **NC** employees were, on average, paid a significantly higher wage (32.2 points) compared to **PC** (22.2 points, $p < .01$, Wilcoxon signed rank test, two-sided, row (9), columns (V) and (VI), Table 3). It is noteworthy that on average **PC** receive wages considerably different from zero.

Observation W1²²: *PC employees receive significantly lower wages than NC.*

Figure 4 gives an overview of average indicated wages depending on a specific employee's profile and a view on the distribution of wages according to employees' general status. It becomes clear that **PC** employees on average receive considerably lower wages compared to **NC** and that most wages they are paid are settled in the low wage range with a decreasing frequency in higher wage margins. In addition, like in ranks, a clear gap between **NC** and **PC** employees can be observed.

²¹A separated conjoint investigation for **male** and **female** employers delivers similar results.

²²W denotes 'Wages'.

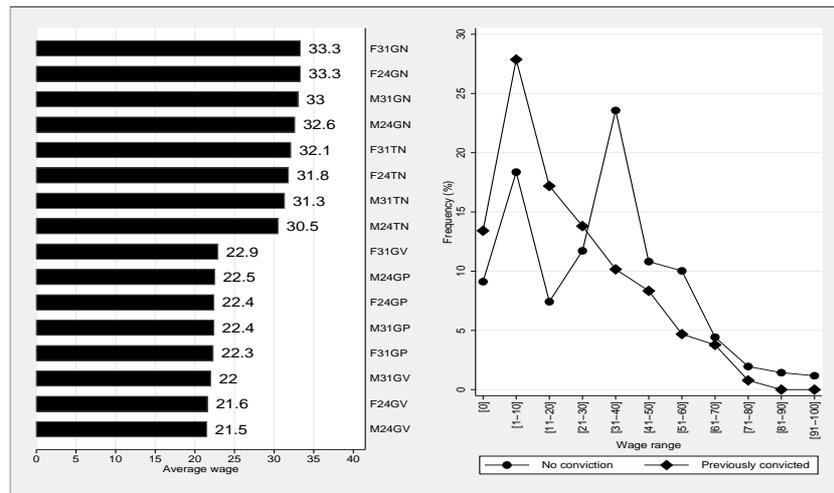


Figure 4: Average and distribution (as categorized according to wage margins) of paid wages.

Our previous finding again allows us to separately analyze both subgroups. Among **NC** employees (column (V), Table 3) participants substantially distinguished between **G** (33.1 points) and **T** employees (32.3 points, $p < .01$, rows (5) and (6)). **Females** (32.7 points) received weakly significantly higher wage payments compared to **males** (31.8 points, $p < .10$, rows (1) and (2)). Between the two age groups (**24** and **31 years old**), no difference in wage payment behavior can be detected (32.0 and 32.5 points, $p = .998$, rows (3) and (4)) (all Wilcoxon signed rank test, two-sided).

Observation W2: Among **NC T** employees are paid lower wages compared to **G**. **Females** receive higher wages than **males**. Employers do not discriminate according to **NC** applicants' age.

Among **PC** employees (column (VI), Table 3), no significant discrimination between property crime (22.4 points) and violence crime (22.0 points) can be detected ($p = .268$, rows (7) and (8)). Moreover, there is no evidence that **female PC** (22.3 points) were significantly favored over **male** ones (22.1, $p = .630$, rows (1) and (2)). Furthermore, both age groups (**24** and **31 years old**) on average received almost similar wage payments (22.0 and 22.4 points, $p = .173$, rows (3) and (4)) (all Wilcoxon signed rank test, two-sided)²³.

Observation W3: Among **PC** no discrimination by employers dependent on type of conviction, sex, or age can be detected.

²³A Tobit regression analysis for paid wages underlines the highly significant and negative impact of STAT and CIT on wage payment. In addition, it can be seen that foreign employers and those participants who expect a high return on investment pay lower wages (see Table 6 in Appendix B).

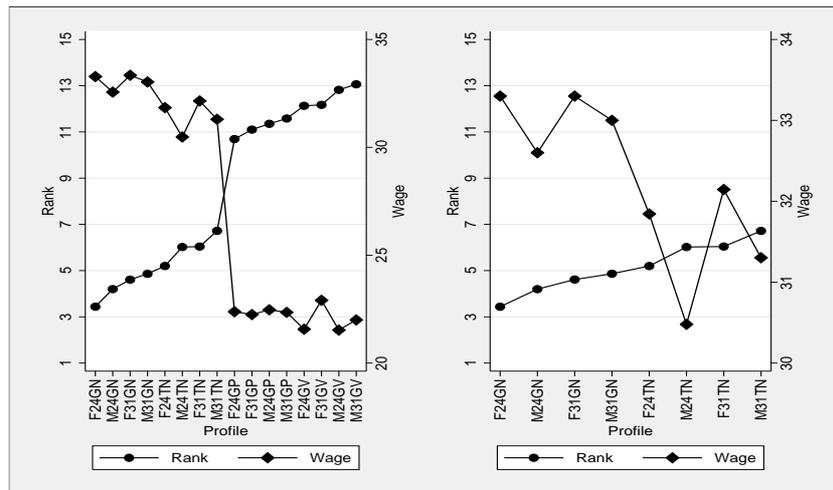


Figure 5: Correlation of ranks and wages.

When we directly compare wages paid to **NC** and **PC** (**G** and **T**; **males** and **females**; **24** and **31 year olds**) we find a highly significant positive correlation (Spearman's $\rho = .889$ ($\rho = .995$; $\rho = .984$; $\rho = .950$), all $p < .01$)²⁴. This indicates that employers hold an individual wage payment attitude *independent* of the treatment²⁵. Even when a specific **PC** (**T**; **male**; **24 year old**) receives a significantly lower wage from her employer compared to another **NC** (**G**; **female**; **31 year old**) who also receives a wage payment from this employer her wage may be considerably higher compared to wages other **PC** (**T**; **males**; **24 year olds**) received from other employers *and*, even more striking, compared to other **NC** (**G**; **females**; **31 year olds**).

Observation W4: *Wage discrimination takes place with regard to individual reference points not according to a common norm.*

In the last step of our analysis on wages we directly contrast assigned average ranks and paid average wages to **NC** and **PC** (**G** and **T**; **males** and **females**; **24** and **31 year olds**). This comparison yields a highly significant negative correlation between ranks and wages for **NC** and **PC** (**males** and **females**; **24** and **31 year olds**) (Spearman's $\rho = -.223$ ($\rho = -.224$; $\rho = -.222$), all $p < .01$). For **G** and **T** we also find a negative correlation. However, it is not statistically significant (Spearman's $\rho = -.420$, $p = .564$) (see Figure 5).

Observation W5: *PC (male; 31 year old) applicants who are less preferred for employment also suffer from lower income when employed. Although 31 year old T applicants are*

²⁴All correlations are carried out two-sided.

²⁵The individual differences between wages paid to **PC** and **NC** further suggest that employers do not consider the relative relation between those wages. Contrary, about 70% of employers decrease **PC**'s income by subtracting a fixed amount between 1 and 10 points compared to wages paid to **NC**.

less preferred in the ranking than **24 year olds** they receive higher wages as employees.

Our findings on wage payment can be summarized in the following result for wages:

Result 2: *Previously convicted employees receive considerably lower wages compared to not previously convicted employees. Within the group of not previously convicted employees highest wages are paid to German females. Turkish males receive lowest wages. Contrary, within previously convicted employees we find that conviction type, sex, or age do not significantly influence employers wage payment decisions. Moreover, employers discriminate with regard to an individual - not universally valid - reference point. In addition, we find previously convicted employees, males, 31 year olds, who are less preferred for employment also to suffer from lower income.*

4.3 Exerted efforts (reciprocity)

In this subsection we report efforts actually exerted by employed second movers dependent on employers' wage payments. For this analysis we averaged all efforts from **NC (PC)**, **G (T)**, **males (females)**, and **24 (31) year old** employees dependent on the wage received from the employer. Figure 6 shows exerted efforts for all types of employees. It is evident that, on average, all groups show higher reciprocity when wages increase (Spearman's $\rho \geq .954$, all $p < .01$).

Observation E1²⁶: *NC and PC, G and T, males and females, 24 and 31 year olds exert higher efforts when receiving higher wages.*

When we compare average exerted efforts dependent on received wages across **NC** and **PC** we find **PC** efforts to be substantially higher compared to **NC** (see Figure 6). Differences in wage margins [1-10] to [61-70] are significant at a $p < .05\%$ - significance level. In addition, **males** exert higher efforts compared to **females** when receiving a wage between 51 and 100 points (all $p < .10$). However, when separately comparing **G** and **T** and **24** and **31 year old** employees we find no statistically significant differences (all $p \geq .426$, all Fisher-Pitman permutation test for two independent samples, two-sided).

Observation E2: *PC exert substantially higher efforts compared to NC when receiving comparable wage payments. G and T as well as 24 and 31 year olds do not exert different efforts. Males are more reciprocal than females when receiving high wages (wages > 50 points).*

²⁶E denotes 'Efforts'.

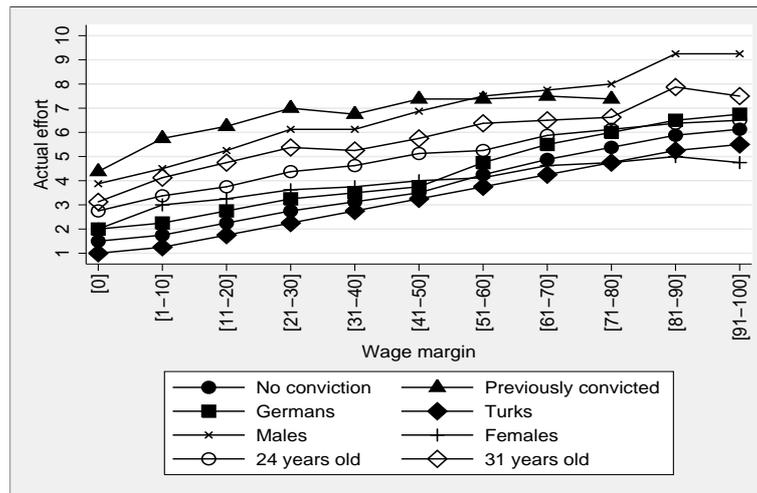


Figure 6: Actual effort of employed PC and NC second movers.

We summarize these findings in the following result for efforts:

Result 3: *All groups of employees exert higher efforts when receiving higher wages. Previously convicted employees and males exert substantially higher efforts compared to not previously convicted employees and females when receiving the same wage (when receiving high wages). Germans and Turks, as well as 24 and 31 year old employees do not exert different efforts.*

4.4 Do specific groups of employees face discrimination?

The evidence found in wage payments leads to the conclusion that the first part of our working hypothesis cannot be rejected. Hence, it is shown that **PC** are paid considerably lower wages and are less preferred for employment. To further investigate motives for employers' discrimination behavior - are **PC** (**T**, **males**, or **31 years olds**) expected to be less trustworthy? -, we now first take a closer look at elicited beliefs on exerted effort. In addition, we will analyze the interplay of stated beliefs and prior hiring preferences and wage allocations at an individual level and under different wage payment systems.

The expected working effort indicates the employers' expectation on the exerted effort which an applicant would contribute in case of employment depending on the wage she is paid. Table 3 and a two-sided Wilcoxon signed rank test show that **PC** employees are expected to exert a lower effort (3.31) compared to **NC** (4.78, $p < .01$, row (9), columns (VII) and (VIII)). Yet, it is reasonable to assume that expected efforts assigned to **PC** employees' are biased downwards because they receive substantially lower wage payments²⁷.

²⁷The same (reversed) argument holds for **NC** employees.

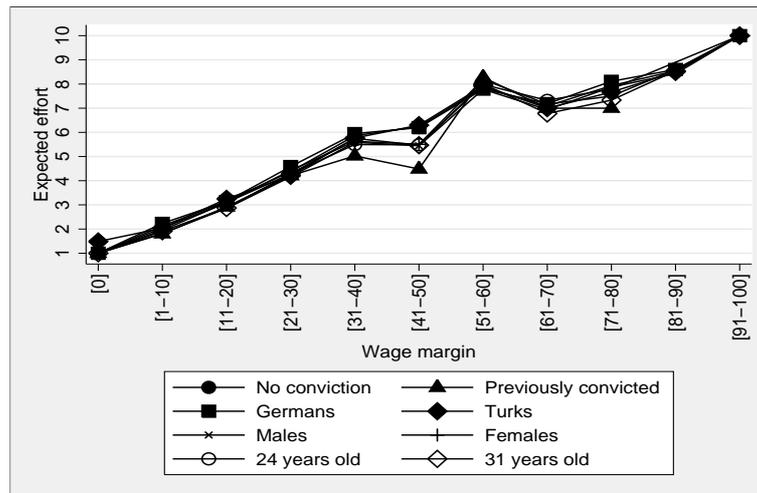


Figure 7: Expected effort dependent on wage margin.

Opposing average wages and corresponding beliefs on exerted efforts yields a strong positive correlation (Spearman's $\rho = .836, p < .01$). To overcome this issue in a first step we look at the different wage margins and corresponding beliefs about efforts separately (see Figure 7). This analysis evidences that in the wage margins [1-10]²⁸ ($p < .05$), [31-40] ($p < .01$), [41-50] ($p < .01$) a significantly higher work effort was expected from **NC** compared to **PC**. However, on the other hand, for the wage margins [0] ($p = 1.000$), [11-20] ($p = .254$), [21-30] ($p = .431$), [51-60] ($p = .268$), [61-70] ($p = .658$), and [71-80] ($p = .125$), all Wilcoxon signed rank test, two-sided), and therefore in 54% of all cases of positive wage payments for **PC**, no substantial differences in employers' expectations are verifiable²⁹. In the most relevant wage margin for **PC** employees ([0-50], see Figure 4), we observe some lower expectations of work effort exerted by them. This, to a certain extent, justifies why **PC** applicants/employees are assigned lower ranks and wages and it mirrors a possible lack of trust toward **PC** trustworthiness. Nevertheless, the fact that we find no consistent pattern of this behavior is an indicator that employers' discriminate independently of an (existing) underlying belief. Employers might penalize **PC** applicants although they do not necessarily expect a comparably lower work effort by them in all cases. A conduct of the same analysis for **G** and **T**, **males** and **females**, and **24** and **31 year olds** yields practically no statistically significant differences³⁰.

Observation D1³¹: *PC employees are believed to exert a lower effort than NC. However,*

²⁸This is the most prominent wage margin for **PC**. For the **NC** group it represents the second most important wage margin.

²⁹Wage margins [81-90] and [91-100] are not considered since no **PC** received a positive wage within these two wage margins (see also Figure 7).

³⁰**G** and **T**: for all wage margins we find $p \geq .212$ (except [0], where $p < .01$); **Males** and **females**: for all wage margins we find $p \geq .154$; **24** and **31 year olds**: for all wage margins we find $p \geq .367$ (except [71-80], where $p < .05$ and [81-90] where no comparison is possible), all Wilcoxon signed rank test, two-sided.

³¹D denotes 'Discrimination'.

given a comparable wage, they are not always expected to exert less effort compared to **NC**. **T**, **females**, and **31 year olds** are not expected to exert lower efforts.

In order to get a standardized and more comparable measure for employers' expectations we next calculate the expected return on investment (eROI). This measure describes the expected outcome for a specific employer given his individual wage payment. It is formally defined as: $eROI = \frac{1}{w} \cdot (100 \cdot e - w)$. Comparing average eROI for **PC** (23.36 fold) with eROI for **NC** (22.26 fold) we find a statistically significant difference ($p < .10$, Wilcoxon signed rank test, two-sided)³². For **T**, **females**, and **31 year olds** we can not find any differences compared to **G** ($p = .744$), **males** ($p = .636$), and **24 year olds** ($p = .241$, all Wilcoxon signed rank test, two-sided).

Observation D2: *Employers expect higher relative returns from **PC** compared to **NC**. **T**, **females**, and **31 year olds** are not expected to generate different eROI compared to **G**, **males** and **24 year olds**.*

When we correlate employers' eROI toward **NC** (**G**; **males**; **24 year olds**) and **PC** (**T**; **females**; **31 year olds**), we find a highly significant positive correlation (Spearman's $\rho = .932$ ($\rho = .965$; $\rho = .962$; $\rho = .979$), all $p < .01$). This shows that employers do not hold a common absolute belief about **PC**'s and **NC**'s (**G** and **T**; **males** and **females**; **24** and **31 year olds**) respective reciprocity levels. If one employer expects a high effort from a **NC** given his wage payment his expectation against **PC** is also relatively high. This might explain our finding from Observation W4 where we show that wages paid to **NC** (**G**; **males**; **24 year olds**) are correlated to wages paid to **PC** (**T**; **females**; **31 year olds**).

Observation D3: *Employers' eROI toward **NC** (**G**; **males**; **24 year olds**) and **PC** (**T**; **females**; **31 year olds**) are highly correlated.*

We can summarize our findings on employers' expectations in the following result:

Result 4: *Ex-offenders are not consistently expected to exert lower efforts compared to not previously convicted job seekers. Contrary to **Turks**, **females**, as well as **31 year olds**, who are not believed to generate lower returns than **Germans**, **males**, and **24 year olds**, employers do even expect higher returns from previously convicted employees. Like wages, employers' eROI toward not previously convicted job seekers, **Germans**, **males**, as well as **24 year olds** on the*

³²Note that in cases where wages are all zero for at least one group, no meaningful values can be calculated for eROI. These observations are therefore neglected (8.33% [=8 subjects] of our observations contain wages all equal to zero)

one hand, and ex-offenders, Turks, females, and 31 year olds on the other hand, are highly correlated.

To gain a better insight *why* **PC** (**T** and **males**) face lower employment opportunities and substantial wage losses we will now analyze the interplay of elicited beliefs and hiring preferences as well as single wage allocations at an individual level. For this analysis we apply the following approach³³: First, we separately investigate each individual's hiring preferences toward **PC** and **NC** by applying a two-sided Fisher-Pitman permutation test for paired replicates. Thus, thereafter we can classify each employer according to whether **PC** were significantly less or more preferred compared to **NC**, or if there is no difference in behavior toward these two groups. In the next step the same procedure is utilized for individual wage payment and eROI using a two-sided Fisher-Pitman permutation test for paired replicates. Combining the individual outcome of the described three-level analysis a behavioral categorization for each of our 96 employers can be generated. Table 4 shows the relevant permutations and the conducted type classification³⁴.

Classification criteria	Hiring	Wage payment	eROI
Type 1: Mistaken Stereotypes	–	–	–
Type 2: Taste based	–	–	o
.	–	–	+
.	+	–	o
.	+	–	+
Type 3: Avoiders	–	o	o
.	–	o	+
.	–	+	o
.	–	+	+
Type 4: No discrimination	o	o	o
Type 5: Positive treatment	+	+	+
.	+	+	o
.	+	o	+
.	o	+	+
.	+	o	o
.	o	+	o
.	o	o	+
Type 6: Others	–/+/o	–/+/o	–/+/o

Table 4: Individual types dependent on hiring, wage payment, and expected return on investment (eROI). A – (+) denotes negative (positive) discrimination. o stands for no statistically significant intra-individual differences.

Applying this individual categorization pattern, can we explain our results with existing theories on discrimination? Economic theories on discrimination can basically be classified into

³³The procedure is exemplarily described for **NC** and **PC**. We use the same approach for analyzing **G** and **T** (**males** and **females**; **24** and **31 year olds**).

³⁴The discriminating classification for ranks, wages and eROI is carried out applying a $p \leq .10\%$ - significance level, two-sided. However, when comparing **G** and **T** the minimal rejection probability of the null-hypothesis is $p = .125$. This is the case when *all* **G** are ranked better than **T**. We therefore then - and only then - also categorize this person as a discriminator.

two main categories: statistical (or rational) discrimination and taste based models. Statistical discrimination describes an income maximizing behavior based on the interpretation of new data in light of prior information (AIGNER and CAIN, 1977; ARROW, 1973; PHELPS, 1972). Contrary, taste based discrimination illustrates a situation wherein individuals are willing to sacrifice money, wages, or profits in order to cater to their prejudice (BECKER, 1957). FERSHTMAN and GNEEZY (2001) have shown that discrimination can also be described as the outcome of these two aspects. They add the possibility of people having stereotypes about others that can be either true or not. In the case that these stereotypes are true, a special case of statistical discrimination arises. When they are not true people have mistaken stereotypes. Even then they may influence the interaction between individuals or groups and lead to biased decisions.

Given our findings on exerted efforts in Observation E2³⁵, within our scheme an employer is supposed to show (1) discrimination because of mistaken stereotypes when this particular employer consistently discriminates in hiring (-) and wage payment (-) (see Table 4). Moreover, this employer believes **PC** to generate a lower eROI (-) than **NC**. Individual taste based discrimination (2) occurs when employers negatively discriminate in wage payment (-) although thinking that **PC** induce a similar or higher ROI compared to **NC** (o or +). This dislike should also be reflected in negatively discriminating hiring decisions (-). However, as taste based discriminators do not maximize monetary income, employers may intentionally want to hire a **PC** (+) to pay them a discriminatory wage afterwards³⁶. In addition to this categorization, a further type of behavior seems plausible: Forming a third group of discriminators, avoiders (3) are antipathetic in hiring **PC**. Nevertheless, once employed they do not negatively discriminate them, neither in wage payment, nor in eROI. The fourth group is represented by those employers who did not display any kind of statistically significant discrimination toward **PC** (4). Another group of interest is given by employers who favor **PC** at least once in hiring, wage payment, or in their expectations on exerted efforts. In group (6) we subsume all employers not fitting the previous classification system.

³⁵**PC(males)** exert substantially higher efforts compared to **NC (females)**; **G** and **T (24 and 31 year olds)** do not differ).

³⁶It is also conceivable that employers do actually believe that **PC** are less trustworthy, however, this underlying belief must not necessarily have influenced their decision. It is conceptually not clear whether this underlying (negative) ex post stated belief ex ante exists *at all*, nor whether it has an impact on employers' decisions (see COSTA-GOMES and WEIZSÄCKER, 2008 for a discussion of this topic).

Classification	NC vs. PC		G vs. T		M vs. F		24 vs. 31	
	w	\bar{w}	w	\bar{w}	w	\bar{w}	w	\bar{w}
(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)
(1) Mistaken Stereotypes	2.08	28.13	0	0	0	0	0	1.04
(2) Taste based	57.29	31.25	0	0	0	0	1.04	0
(3) Avoiders	22.92	17.71	20.83	20.83	1.04	1.04	14.58	14.58
(4) No discrimination	5.21	5.21	69.79	69.79	86.46	81.25	66.67	67.71
(5) Positive treatment	2.08	2.08	1.04	1.04	3.13	6.25	9.38	8.33
(6) Others	10.42	15.63	8.33	8.33	9.38	11.46	8.33	8.33

Table 5: Frequency (in %) of individual employer types dependent on treatment and wage payment system. We denote M=male, F=female, 24=24 years old, and 31=31 years old.

Table 5 shows the distribution of employer types dependent on our classification criteria and treatment variables. At least 85 out of 96 employers (= 88.54%) fall into the classification of Types 1-5, independently from the investigated treatment. For **NC** and **PC** we find that most employers can be categorized as taste based discriminators (57.29% = 55 employers). Only 2 employers (= 2.08%) (wrongly) expected **PC** to actually generate lower ROI than **NC**³⁷. When we look at the third type of our classification system we find 22 employers (= 22.92%) to be avoiders. Moreover, two of them (= 2.08%) favored **PC** at least once. Only 5 employers (= 5.21%) did not show any kind of discrimination (column (II), rows (2-6)). Looking at **G** and **T** (**males** and **females**; **24** and **31 year olds**) we practically find no discrimination because of mistaken stereotypes or based on taste. Most employers did not significantly discriminate against these groups (69.79% (86.46%; 66.67%)) (columns (IV), (VI), and (VIII), row (4)). In case of **G** and **T** (**24** and **31 year olds**) 20.83% (14.58%) tried to avoid an employment of **T** (**31 year olds**) (columns (IV) and (VIII), row (3)). However, another 9.38% (= 9 employers) favored **31 year old** employees over **24 year olds** (column (VIII), row (5)).

Observation D4: *Under an individual wage payment system most employers are taste based discriminators or mere avoiders toward PC compared to NC. Against T, males, 31 year olds no significant discrimination can be detected on the individual level. However, some employers tried to negatively discriminate, others favored 31 year olds.*

The previous analysis was carried out based on employers' individual wages and associated beliefs on reciprocity. To examine the effects of an introduced fixed wage system where the employee is paid a maximum wage, we also asked employers about their expectations on employees' efforts in case a fixed wage $\bar{w} = 100$ points would be paid by them to the employees³⁸. In that case employers believe **NC** (6.08 fold) (**females** (5.65 fold)) to generate higher ROI

³⁷This finding is supported by the fact that wages and eROI are negatively correlated for **NC** (Spearman's $\rho = -.605, p < .01$) and **PC** (Spearman's $\rho = -.560, p < .01$).

³⁸This virtual situation contained no monetary incentives.

compared to **PC** (5.03 fold, $p < .01$) (**males** (5.46 fold, $p < .05$) (all Wilcoxon signed rank test, two-sided). For **T** and **31 year olds** we can not find any statistical differences compared to **G** ($p = .650$) and **24 year olds** ($p = .785$, both Wilcoxon signed rank test, two-sided). When correlating employers' eROI toward a specific treatment group under individual and fixed wage payment we find no significant correlations³⁹. This indicates that employers' eROI under the fixed wage payment system are not sufficiently biased by their eROI under the individual wage payment system and allows us to compare behavior manifested under both systems.

Observation D5: *In a fixed (high) wage payment system employers' eROI are significantly higher toward **NC (females)** compared to **PC (males)**. **T** and **31 year olds** are not expected to generate different ROI compared to **G** and **24 year olds**.*

We now categorize employers according to our classification system given a fixed (high) wage payment system. We find that 28.13% (= 27 employers) of 96 employers hold a mistaken stereotype about **PC's** reciprocity⁴⁰. 31.25% (= 30 employers) are classified as taste based discriminators, 17.71% (= 17 employers) as avoiders. Those 5 (2) employers (= 5.21% (2.08%)) who showed no discrimination (a positive treatment) under individual wages behaved consistently under both wage payment systems (column (III), rows (1)-(6), Table 5)⁴¹. Analyzing **G** and **T (males and females; 24 and 31 year olds)** we again practically find no discrimination because of mistaken stereotypes or based on taste. Toward **G** and **T (males and females; 24 and 31 year olds)** most employers did not significantly discriminate (69.79% (81.25%; 67.71%)) (columns (V), (VII), and (IX), row (4)). In case of **G** and **T (24 and 31 year olds)** 20.83% (14.58%) of employers want to avoid an employment of **T (31 year olds)** (columns (V) and (IX), row (3)). In addition, 8.33% (= 8 employers) favored **31 year old** employees over **24 year olds** (column (IX), row (5)).

Observation D6: *Under a fixed (high) wage system, toward **PC**, most employers are taste based discriminators. However, mistaken stereotype discriminators and avoiders are also representative groups. Toward **T, males, 31 year olds** no significant discrimination on the individual level can be detected. Some employers tried to negatively discriminate, others favored **31 year olds**.*

³⁹For **NC**: Spearman's $\rho = -.069$; **PC**: $\rho = .034$; **G**: $\rho = -.067$; **T**: $\rho = -.047$; **males**: $\rho = -.016$; **females**: $\rho = .011$ **24 year olds**: $\rho = -.001$; **31 year olds**: $\rho = -.019$, all $p \geq .522$.

⁴⁰This finding is based on the assumption that **PC** employees also exert higher efforts than **NC** under the fixed (high) wage payment system. If this assumption does not hold we observe a special case of statistical discrimination. Furthermore, we assume a similar employers' hiring pattern.

⁴¹When looking at individuals we find them to be the same employers.

When we compare the distribution of types under both payment systems two points become apparent: First, while we find not much evidence on discrimination because of mistaken stereotypes against **PC** under individual wages (2.08%), the amount of employers showing this kind of discrimination under fixed wages almost reaches 30%. Second, the percentage of employers showing avoidance, no discrimination, or favoring of a specific treatment group remains notably stable across wage payment systems.

Observation D7: *Under a fixed (high) wage system, more employers, compared to individual wages, discriminate against **PC** because they believe them to be less trustworthy than **NC**. The percentage of employers showing avoidance, no discrimination, or favoring of a specific treatment group remains notably stable across wage payment systems.*

In order to examine whether the nature of discrimination under individual wages is dependent on the potential target of discrimination, in a last step we look at the distribution of employer types given distinct treatment groups. It becomes evident that while discriminatory behavior against **PC** is very consistently found in hiring and/or wage payment, we find a disadvantage of **T** and **31 year olds**, if detectable, to be more a phenomenon of avoidance. Moreover, gender discrimination is hardly identifiable on the individual level.

Observation D8: *Discrimination against **PC** is basically taste based whereas the disadvantage of **T** and **31 year olds** mainly consists of avoidance. Gender discrimination is hardly identifiable on an individual level.*

Now we can summarize our findings on discrimination in the following result:

Result 5: *Under the given individual wage payment system most employers are taste based discriminators or mere avoiders. In a fixed (high) wage payment system employers' eROI are significantly higher toward not previously convicted employees compared to ex-offenders. Turks, females, 31 year olds are not believed to generate lower returns than Germans, males, 24 year olds. Toward ex-offenders, most employers are taste based discriminators but mistaken stereotype discriminators and avoiders are also representative groups. Under both payment systems Turks, males, and 31 year olds are not significantly discriminated on an individual level. Moreover, some employers tried to avoid, others were looking for an employment of 31 year olds. When comparing the distribution of employer types under both payment systems we find that under a fixed wage system more employers discriminate against previously convicted employees because they believe them to be less trustworthy. In addition, the percentage of employers showing either no discrimination or favoring of a specific treatment group remains notably stable. Finally, examining the nature of discrimination dependent on*

the potential target of discrimination, we find discrimination against ex-offenders to be basically taste based. Contrary, the disadvantage of Turks and 31 year olds mainly consists of avoidance. Gender discrimination is hardly identifiable on an individual level.

5 Conclusion

In this chapter we experimentally investigated whether job seekers or employees holding a criminal record are less trusted and expected to be less trustworthy compared to not previously convicted workers. Conducting a controlled gift exchange experiment with criminal offenders acting as job applicants/employees and students deciding as employers we were able to systematically and directly study the impact of an existent previous conviction on employers' trust behavior and expectations toward different groups of applicants/employees. In addition, our experimental setup also allowed us to compare employers' discriminatory behavior against ex-offenders with their attitudes against other potential target groups of discrimination as foreigners, women, or elder workers.

Our results give substantial support for a disadvantage of previously convicted and Turkish employees who are less preferred for employment and paid significantly lower wages compared to not previously convicted and German workers. However, we find associated beliefs about convicted and Turkish employees' reciprocity not to play a primary role in interactions with them. Only under a fixed wage system, where employers must fully trust employees, ex-offenders are believed to be less trustworthy. Thus, we can show that the reluctance to hire an ex-offender or Turk only partly consists of employers' low expectations concerning their trustworthiness. Analyzing employers' behavior on an individual level we find that employers' discrimination against ex-felons is mainly taste based or avoidance driven. Although, on average, we find a strong negative discrimination of Turkish workers, at the individual level they are basically avoided by employers and not wage discriminated. For females and males the evidence is mixed: Employers preferred to hire females regardless of whether they have an empty record or not. Although not previously convicted females also received higher wage payments compared to males, sex played no significant role in wage payments toward ex-offenders. Age, on average, had practically no influence on employers' decisions, although employers preferred to hire younger applicants compared to elder. On the individual level we can see that several employers tried to negatively discriminate 31 year old employees while others favored them.

We have furthermore shown that employers' reluctance to hire ex-offenders and to pay them an appropriate wage is considerably stronger than their aversion to hire Turkish, male, or elder

applicants. In addition, we find that employers discriminate specific target groups differently. Whereas discrimination against ex-offenders is basically taste based and detectable both on the individual and the aggregate level, the disadvantage of Turkish or elder workers mainly consists of individual avoidance, not of income discrimination⁴². However, on the aggregate level Turkish employees also receive significantly lower wages compared to German workers.

Another contribution of this chapter is the finding that wage discrimination occurs across all target groups according to individual reference points and not because of a commonly known or valid standard. This implies that ex-offenders or Turks, although individually discriminated, are not always less trusted (do not always receive utterly low wages). Compared to other employees - previously convicted or not - their wages might be even higher.

Given these findings, the fact that in our study previously convicted workers are more trustworthy and because we provided employers with no additional information about applicants/employees, we can conclude that previously convicted workers suffer from stigmatization rather than from negative expectations concerning their trustworthiness. This replicates and extends the findings of HOLZER et al. (2002), WALDFOGEL (1994a), BECKER (1968), and SCHWARTZ and SKOLNICK (1962). Nevertheless, it is thinkable, that employers hold negative beliefs about ex-offenders independently of the information provided through personal profiles in our experiment and besides considerations concerning actual trustworthiness. It is possible that employers are inferring social background from the existence of a conviction. In the extreme form, employers do not care at all about conviction but are discriminating only against the associated social background conveyed by the fact of the existence of a conviction. Contrary, Turkish workers are mainly discriminated by avoiding them. How can we explain this difference in discrimination patterns? One possible reason for this difference might lay in the perception and awareness of different social groups. Although ex-offenders as shown in the beginning represent a numerically large group they are mostly invisible - a criminal record is not easily detectable and no one has an incentive to voluntarily talk about one's own crime history. Furthermore, usually only experts or social workers directly get in touch with criminal offenders on purpose. Their problems and needs are not in public focus. Contrary, foreigners are usually easily recognizable, either by their appearance or by their language expression. Moreover, there is normally no reason to hide one's own origin. Beyond that, race discrimination is very much in the focus of public discussion and most people are sensitized to

⁴²Our results on employees' efforts must be handled with some care: The number of observations is limited to 8. Moreover, it is possible that (previously convicted) employees behaved somewhat socially desirable or suspiciously, although we explicitly guaranteed anonymity. All of them were totally unexperienced concerning experiments and the game was played only one round. Nevertheless, these facts do not change the finding of mainly taste based discrimination.

this topic. Hence, it seems to be rather important to act politically correct toward foreigners (a random treatment group). Against former criminals (not a pure random treatment group) this argument appears to be reversed - it might be even socially desirable to punish those who have violated the law and harmed the society.

The preferential treatment of (not previously convicted) female workers may indicate that women are more trusted and expected to be more trustworthy compared to men. Since we were able to eliminate any entrance barriers typically existing for women in some branches such as hard physical or endurance based work (e.g., production line work, mechanic work), it can be shown that women even have an advantage over men.

The results of our study advise to invest in an increased preparatory and accurate qualification of prison inmates as well as to take care of them enduringly after release in order to support professional and social rehabilitation by means of effective education and, most importantly, its certification as well as signalling it to the public. Furthermore, with the help of extensive clarification of their actual post release behavior, negative and (not trustworthiness-related) discriminating expectations and lack of trust as well as selective employment decisions which results consequentially have to be dispelled among potential employers. Incentives for an enlarged employment of former prison inmates and a more intensive engagement for this group of employers have to be established. As we have found that Turkish workers are mainly avoided by employers and discriminated on the aggregate level, although they are not expected to be less trustworthy, we suggest to increasingly care for immigration programs to bring foreigners and natives together, e.g., in working teams, to bridge invisible lines of segregation within society.

For future work we suggest a systematical and separate investigation of employer's ex ante elicited complex expectations on employees' behavior. This is to not only study beliefs associated with trustworthiness but also with traits and characteristics only indirectly connected to reciprocity. In addition, employers' hiring behavior can be in the focus on interest when fully enforceable contracts are given. Moreover, the effect of previous baseless impeachment and its stigmatizing impact on employment opportunities and income levels - of different social groups (e.g., foreigners) - are interesting approaches to plot a more differentiated picture of employers' behavior and to draw valid conclusions about ex-offenders labor market performance.

References

- AIGNER, D. and G.G. CAIN (1977): "Statistical Theories of Discrimination in Labor Markets", *Industrial and Labor Relations Review*, 30(2), 175-187.
- AKERLOF, G.A. (1982): "Labor Contracts as Partial Gift Exchange", *The Quarterly Journal of Economics*, 97(4), 543-569.
- ARROW, K. (1973): "The Theory of Discrimination", in: O. ASHENFELTER and A. REES (Eds.), *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press, pp. 3-33.
- BACKHAUS, K., B. ERICHSON, W. PLINKE, and R. WEIBER (1994): *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung. Siebte, vollständig überarbeitete und erweiterte Auflage.* Berlin u.a.: Springer Verlag.
- BECKER, G.S. (1961): *The economics of discrimination.* Second Edition, Chicago: University of Chicago Press, 1961.
- BECKER, G.S. (1968): "Crime and Punishment: An Economic Approach", *The Journal of Political Economy*, 76(2), 169-217.
- BENNA, R. (1998): *Bedarfsorientiertes Filialbanking. Empirische Identifikation erforderlicher Leistungsstrukturen mit Hilfe der Conjoint-Analyse.* Frankfurt am Main: Fritz Knapp Verlag (Schriftenreihe des Zentrums für Ertragsorientiertes Bankmanagement, Bd. 16).
- BERTRAND, M. and S. MULLAINATHAN (2004): "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination", *American Economic Review*, 94(4), 991-1013.
- BOSHIER, R. and D. JOHNSON (1974): "Does Conviction Affect Employment Opportunities?", *British Journal of Criminology*, 14(3), 264-268.
- BUIKHUISEN, W. and F.H.P. DIJKSTERHUIS (1971): "Delinquency and Stigmatization", *British Journal of Criminology*, 11, 185-187.
- BUONANNO, P. (2003): "The Socioeconomic Determinants of Crime. A Review of the Literature", Working Papers 63, University of Milano-Bicocca, Department of Economics, revised Nov. 2003.
- CASPI, A., B.R.E. WRIGHT, T.E. MOFFIT, and P.A. SILVA (1998): "Early Failure in the Labor Market: Childhood and Adolescent Predictors of Unemployment in the Transition to Adulthood", *American Sociological Review*, 63(3), 424-451.
- COOPER, D.J., J.H. KAGEL, W. LO, and X.L. GU (1999): "Gaming Against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers", *American Economic Review*, 89(4), 781-804.

- COSTA-GOMEZ, M.A. and G. WEIZSÄCKER (2008): "Stated Beliefs and Play in Normal-Form Games", *Review of Economic Studies*, 75(3), 729-762.
- DARITY, W.A. Jr. and P.L. MASON (1998): "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender", *The Journal of Economic Perspectives*, Vol. 12 (2), 63-90.
- FALK, A. and E. FEHR (2003): "Why labour market experiments?", *Labour Economics*, 10(4), 399-406.
- FEHR, E., G. KIRCHSTEIGER, and A. RIEDL (1998): "Gift exchange and reciprocity in competitive experimental markets", *European Economic Review*, 42(1), 1-34.
- FEHR, E. and J. LIST (2004): "The hidden costs and returns of incentives - trust and trustworthiness among CEOs", *Journal of the European Economic Association*, 2(5), 743-771.
- FERSHTMAN, C. and U. GNEEZY (2001): "Discrimination in a segmented society: An experimental approach", *The Quarterly Journal of Economics*, 116(1), 351-377.
- FREEMAN, R.B. (1992): "Crime and the Employment of Disadvantaged Youth", in: G. PETERSON and W. VROMAN (Eds.), *Urban Labor Markets and Job Opportunity*, Washington, DC: Urban Institute Press, pp. 201-37.
- GOLDBERG, A., D. MOURINHO, AND U. KULKE (1996): "Labour market discrimination against foreign workers in Germany", *International Migration Papers*.
- GRANOVETTER, M. (1995): *Getting a Job: A study of Contracts and Careers*. Second Edition, Chicago: University of Chicago Press.
- GREEN, P.E., D.S. TULL, and G. ALBAUM (1988): *Research for Marketing Decisions*. Fifth Edition, Englewood Cliffs, New Jersey: Prentice Hall.
- GROGGER, J. (1995): "The Effect of Arrest on the Employment and Earnings of Young Men", *The Quarterly Journal of Economics*, 110(1), 51-71.
- HAIR, J.F., R.E. ANDERSON, R.L. TATHAM, and W.C. BLACK (1995): *Multivariate Data Analysis with Readings*. Fourth Edition, Englewood Cliffs, New Jersey: Prentice Hall.
- HOLZER, H.J., S. RAPHAEL, and M.A. STOLL (2001): "Will Employers hire Ex-Offenders? Employer Preferences, Background Checks, and Their Determinants", in: M. PATILLO-McCOY, D. WEIMAN, and B. WESTERN (Eds.), *The Impact of Incarceration on Families and Communities*, New York: Russell Sage Foundation.
- HOLZER, H.J., S. RAPHAEL, and M.A. STOLL (2002): "Can Employers Play a More Positive Role in Prisoner Reentry?", Working Discussion Paper for the Urban Institute's Reentry Roundtable: Can Employers Play a More Positive Role in Prisoner Reentry?

- KLING, J.R. (1999): "The Effect of Prison Sentence Length on the Subsequent Employment and Earnings of Criminal Defendants", *Woodrow Wilson School Discussion Papers in Economics*, No. 208. Princeton, NJ, Princeton University.
- LIYANARACHCHI, G.A. (2007): "Feasibility of using student subjects in accounting experiments: a review", *Pacific Accounting Review*, 19(1), 47-67.
- LOTT, J.R. (1990): "The Effect of Conviction on the Legitimate Income of Criminals", *Economics Letters*, 34, 381-385.
- LYNCH, J. and W. SABOL (2001): "Prisoner reentry in perspective", *Crime Policy Report*, Vol. 3, Washington, DC: The Urban Institute.
- MOFFIT, T.E. (1993): "Life-course-persistent and adolescence-limited antisocial behavior: A developmental taxonomy", *Psychological Review*, 100, 674-701.
- NEEDELS, K.E. (1996): "Go Directly to Jail and Do Not Collect? A Long-Term Study of Recidivism, Employment, and Earnings Patterns among Prison Releases", *Journal of Research in Crime and Delinquency*, 33(4), 471-496.
- PAGER, D. (2003): "The Mark of a Criminal Record", *American Journal of Sociology*, 108(5), 937-975.
- PHELPS, E.S. (1972): "The Statistical Theory of Racism and Sexism", *American Economic Review*, 62(4), 659-61.
- ROSCIGNO, V.J., S. MONG, R. BYRON, and G. TESTER (2007): "Age Discrimination, Social Closure and Employment", *Social Forces*, 86(1), 313-334.
- ROSE, D. and T. CLEAR (1998): "Incarceration, social capital, and crime: implications for social disorganization theory", *Criminology*, 36, 441-480.
- SABOL, W.J. and J.P. LYNCH (1998): "Assessing the Longer-run Consequences of Incarceration: Effects on Families and Employment", Conference Paper, New York.
- SANCHEZ-JANKOWSKI, M. (1991): "Island in the street: Gangs and American urban society", University of California Press, Berkeley.
- SCHWARTZ, R.D. and J.H. SKOLNICK (1962): "Two Studies of Legal Stigma", *Social Problems*, 10, 133-142.
- SULLIVAN, M.L. (1989): *Getting paid: Youth crime and work in the inner city*. Ithaca, NY: Cornell University Press.
- UGGEN, C. and J. MANZA (2002): "Democratic Contraction? The Political Consequences of Felon Disenfranchisement in the United States", *American Sociological Review* 67(6), 777-803.
- VENKATESH, S.A. (2000): *American project: The rise and fall of a modern Ghetto*. Cambridge, MA: Harvard University Press.

- WALDFOGEL, J. (1994a): "The Effect of Criminal Conviction on Income and Trust Reposed in Workmen", *Journal of Human Resources*, 29, 62-81.
- WALDFOGEL, J. (1994b): "Does Conviction Have a Persistent Effect on Income and Employment?", *International Review of Law and Economics*, 14, 103-119.
- WEICHSELBAUMER, D. (2004): "Is it Sex or Personality? The Impact of Sex-Stereotypes on Discrimination in Applicant Selection", *Eastern Economic Journal*, 30(2), 159-86.
- WESTERN, B. and K. BECKETT (1999): "How unregulated is the U.S. labor market? The U.S. penal system as a labor market institution", *American Journal of Sociology*, 104(4), 1030-1060.
- WESTERN, B., J.R. KLING and D.F. WEIMAN (2001): "The Labor Market Consequences of Incarceration", *Crime & Delinquency*, 47(3), 410-427.
- WESTERN, B. (2004): "The Effects of Incarceration on Wages and Employment", Department of Sociology, Princeton University.

Appendix A: Incentive compatible matching mechanism⁴³

Consider the following allocation game, henceforth denoted by AG . There are m objects which shall be allocated among n agents; all agents receive at most one object (potentially less if $m < n$). The players' preferences over the objects \succeq_i are represented by utility functions

$$u_i : \mathcal{M} \rightarrow \mathbb{R}, \quad i = 1, \dots, n,$$

where \mathcal{M} is the set of objects. The allocation procedure is as follows:

1. To begin with, each agent i has to indicate his preferences over the m objects, i.e. he has to state an m -tuple $s_i = (s_i^1, \dots, s_i^m)$ such that $s_i^l \in \mathcal{M}$, $l = 1, \dots, m$, and $s_i^l \neq s_i^k$ if $l \neq k$. Accordingly, for each agent i , the strategy space S_i of this agent can be represented by the set P^m of all permutations of $\{1, \dots, m\}$, i.e. $S_i = P^m$ for all i .
2. Once all agents have determined their strategy, the order in which the n agents are supplied is determined by means of a random device ρ . In particular, ρ determines an n -tuple $p^n = (p_1, \dots, p_n)$ with $p_k \in \{1, \dots, n\}$, $k = 1, \dots, n$, and $p_k \neq p_l$ if $k \neq l$ so that agent p_1 is supplied first and agent p_n is supplied last; note that again $p^n \in P^n$ where P^n is the set of all permutations of $\{1, \dots, n\}$. The probability that p^n is chosen is denoted by $\rho(p^n)$. By assumption $\rho(p^n) > 0$ for all $p^n \in P^n$.
3. After a certain supply order $p^n = (p_1, \dots, p_n)$ has been determined, the objects are allocated among the agents so that, once it is his turn to be supplied, each agent i receives that object out of the remaining ones which is ranked best according to s_i . In particular, first, agent p_1 is supplied and receives the object $s_{p_1}^1$. Then, agent p_2 is supplied and receives the object $s_{p_2}^k$ where $k := \min\{\kappa \in \{1, \dots, m\} \mid s_{p_2}^\kappa \in \mathcal{M} \setminus \{s_{p_1}^1\}\}$, and so forth until either all objects contained in \mathcal{M} are allocated or else each agent has received one object.

Finally, for all i , agent i 's ex ante expected utility for given strategy profile $s = (s_1, \dots, s_n)$ is denoted by

$$U_i(s; \rho) = \sum_{p \in P^n} \rho(p) u_i(o(s; p)),$$

where $o(s, p)$ denotes the object that is allocated to agent i according to the supply order p and given s .

Lemma 1 *Consider the allocation game AG defined above. For each agent i , let S_i^* be the set of all strategies s_i^* which represent agent i 's true preferences, i.e. $s_i^{*1} \succeq_i \dots \succeq_i s_i^{*m}$ for*

⁴³The formal proof presented here is taken from WICHARDT (2008) and is reproduced here with the consent of the author; see P. C. WICHARDT (2008): "Allocating few objects on many agents in lab experiments", mimeo, Department of Economics, University of Bonn.

all $s_i^* \in S_i^*$ (note that S_i^* is not a singleton set if agent i is indifferent between some objects). Then, for each agent i , it holds that:

1. for all $s_{-i} \in S_{-i}$ and any supply order $p \in P^n$, $S_i^* \subset \operatorname{argmax}_{s_i} u_i(o(s_i, s_{-i}; p))$, i.e. for any given strategies s_{-i} of the other agents and any supply order p , each strategy $s_i^* \in S_i^*$ maximizes agent i 's utility; hence, for all $s_{-i} \in S_{-i}$ $S_i^* \subset \operatorname{argmax}_{s_i} U_i(s_i, s_{-i}; \rho)$, i.e. irrespective of the other agents strategy choices, stating the true preferences maximizes agent i 's expected utility.
2. if $m \leq n+1$, then the strategies $s_i^* \in S_i^*$ weakly dominate all other strategies $s_i \in S_i \setminus S_i^*$.
3. if $m > n+1$, then for all $s_{-i} \in S_{-i}$ and any supply order $p \in P^n$, $\tilde{S}_i^* \subset \operatorname{argmax}_{s_i} u_i(o(s_i, s_{-i}; p))$, where \tilde{S}_i^* is the set of all strategies \tilde{s}_i^* with $\tilde{s}_i^{*1} \succeq_i \dots \succeq_i \tilde{s}_i^{*n}$, and the strategies $\tilde{s}_i^* \in \tilde{S}_i^*$ weakly dominate all strategies $s_i \in S_i \setminus \tilde{S}_i^*$.

Proof. The proof of Lemma 1 is rather intuitive. Statement 1, in fact, is immediate. If agent i chooses $s_i^* \in S_i^*$, then for any given supply order p agent i gets that object out of the remaining ones which is ranked best according to s_i^* . Since s_i^* reflects agent i 's true preferences the respective object, $o(s_i^*, s_{-i}; p)$, maximizes u_i with respect to the set of remaining objects, i.e. $s_i^* \in \operatorname{argmax}_{s_i} u_i(o(s_i, s_{-i}; p))$ (this is a consequence of the independence axiom). The second part of Statement 1 is a direct consequence of the first.

In order to prove Statement 2, we show that for any strategy $s_i \in S_i \setminus S_i^*$ there is a combination of strategies of the other agents s_{-i} such that s_i gives a strictly inferior payoff to any $s_i^* \in S_i^*$ for at least one possible supply order p . Since $s_i^* \in S_i^*$ maximizes u_i for all possible supply orders and all $s_{-i} \in S_{-i}$ (Statement 1) and because of the fact, by assumption, all possible supply orders are chosen with positive probability this implies that s_i is weakly dominated by all $s_i^* \in S_i^*$. In particular, consider $s_i \in S_i \setminus S_i^*$ and let $k < n+1$ be the smallest index such that $s_i^k \prec s_i^l$ for some $l > k$ (by definition of s_i such a k exists). Moreover, let p be such that agent i is the k th agent to be supplied and let s_{-i} be such that the set of objects allocated to the agents $j = p_1, \dots, p_{k-1}$ is given by $\{s_i^1, \dots, s_i^{k-1}\}$. Then, it is $o(s_i, s_{-i}; p) = s_i^k$, i.e. when being supplied agent i object s_i^k . Yet, by construction, $s_i^l \succ s_i^k$. Hence, s_i does not maximize u_i given s_{-i} and p . From Statement 1 it, thus, follows that $u_i(s_i, s_{-i}; p) < u_i(s_i^*, s_{-i}; p)$ for all $s_i^* \in S_i^*$, which was to be shown. Finally, the strategies $s_i^* \in S_i^*$ are only weakly dominant, because if s_{-i} is such that the other agents never choose agent i 's most preferred object, say o_i^* , then every strategy s_i with $s_i^1 = o_i^*$ delivers the same payoff for agent i .

The proof of Statement 3 is analogous to the combination of the preceding arguments except that one now has to take into account that differences in s_i after the n th entry are of no effect for agent i 's utility because at most n object are allocated. ■

Corollary 1 Consider the allocation game AG. If some agent i believes that any strategy profile $s_{-i} \in S_{-i}$ is possible, i.e. his beliefs assign positive probability to any s_{-i} being played, then it is a strict best response for this agent to state his true preferences, i.e. to play $s_i^* \in S_i^*$.

Appendix B: Tobit regression analysis for dependent variable 'wage'

Variable	Coefficient	(Standard error)
SEX	.5458698	(.4556196)
AGE	.3934227	(.455307)
CIT	-1.704945	(.6357706)***
STAT	-11.22571	(.558081)***
eROI	-.1533305	(.0155695)***
Employers' sex	.7112371	(1.02929)
Employers' age	.0563941	(.1450921)
Employers' citizenship	-6.613719	(1.77852)***
Employers' no. of siblings	-.0628913	(.4750061)
Size of city employer most lived in	.1223247	(.4777783)
Employers' disposable monthly income	-.000225	(.0026869)
Constant	54.02764	(4.1958)***
Prob > χ^2		.0000
N	1.536	(96 groups)
ρ	.7552951	(.0103187)

Table 6: Coefficients from Tobit regression analysis for dependent variable WAGE PAYMENT. Asterisks indicate variables as being significant at 1% - significance level.

Appendix C: Instructions for employers

Instructions for the Experiment

You are taking part at an economical experiment. Please read the following explanations and instructions carefully. It is important not to speak to any other participant throughout the entire experiment. In case you do not understand anything, please refer to these instructions. If you still have any questions afterwards, raise your hand. We will come to your cabin and answer your question personally.

In order to guarantee your anonymity you will draw a personal number at the beginning of the experiment. With the personal number we are not able to conclude which participant gave a certain decision or answer in the experiment.

At the beginning of the experiment you will receive an endowment of 10 EURO. During the experiment you are able to raise your account by collecting points. The amount of points, which you can gain during the experiment, depends on your and the other participants' decisions. All points will be calculated into EURO at the end according to the following measurement:

$$1 \text{ Point} = 0.10 \text{ EURO}$$

During the experiment you are interacting with other decision makers. A part of the other decision makers will determine their decisions in a separate experiment session after your session. Because of that, we are not able to pay off your earned income today. However, you are allowed pick it up in the office of the Laboratory for Experimental Economics daily between 9am and 12am, starting Monday, August 28th 2006. Alternatively, we can also transfer your total income to your bank account.

At no point in time you will get to know the name of any other decision maker, just as the other decision makers will not get to know your identity.

The experiment is divided into a number of steps. At the end of the experiment we will ask you to answer a set of questions, which we need for statistical evaluation of the experiment. Your answers will be evaluated anonymously.

Experiment-overview

Within this experiment it exists a labor market of 16 employers (he) and 16 applicants (she). Employers and applicants both can earn an income. Each applicant applies at every employer with her personal data sheet in order to be hired for a working project lasting one period. After today's session each employer will be assigned one applicant, depending on the decisions of all employers. The assigned applicant then actually interacts with her employer. Accordingly, after the session there will be 16 employer-applicant-couples. After the applicant is employed, every employer pays her a wage. The employed applicant chooses her effort afterwards.

During the entire experiment you are an employer. All other experimental participants who are joining today's session are also employers.

The experiment consists of three steps. You as an employer have to make decisions in two of the three steps today.

1st step: In this step you create a personal rank order of the applicants.

2nd step: You predefine a wage for each possible applicant before you know which applicant you will be assigned. The wage level can be individually predefined for each applicant.

At this point, today's session will be closed after you have answered a questionnaire. Before you fill in the questionnaire you will get to know which applicant you are going to employ.

3rd step: This step is set up at a different point in time after this experiment, in a separate session. At the beginning, applicants will be informed about the amount of wage the employer, who engages them, is going to pay them. Thereupon, applicants choose their effort. Applicants do not receive any other information about their assigned employer. Accordingly, applicants do not know that you had to choose between several applicants and they also do not know your personal ranking order. Applicants can choose their effort freely. They receive the wage, which their employer has predefined, independently of their effort.

Detailed description of experimental procedure

1st step: The employer's decision of the ranking order

In this step you set up your personal ranking order of the 16 applicants on your computer screen. Therefore, 16 data sheets with information about the 16 applicants will be made available. The applicants' data sheets are set up in the following way. Hereby, 'text' varies of course among the applicants.

Age	<i>text</i>
Sex	<i>text</i>
Citizenship	<i>text</i>
Family status	<i>text</i>
Education	<i>text</i>
Occupational skills	<i>text</i>
Additional professional knowledge	<i>text</i>
Additional remarks	<i>text</i>

Please bear in mind:

- a) Applicants are real people.
- b) Applicants actually hold the qualities named on their data sheets.

It is your task to scan the 16 data sheets and to set up a personal ranking order. Therefore, you need to give each applicant a ranking number between 1 and 16, with rank 1 being the applicant you mostly prefer and with rank 16 being the applicant you would like to employ last. All employers set up their ranking order simultaneously on their computer. With your personal ranking order you influence the probability of employing a certain applicant and the chance to interact with her.

The entry mask on the computer for your decision of your ranking order looks like this:

1. Schritt: Rangordnungsentscheidung

Karrierestufe	Rang
top	<input type="text"/>
ptb	<input type="text"/>
abr	<input type="text"/>
mca	<input type="text"/>
org	<input type="text"/>
nip	<input type="text"/>
joe	<input type="text"/>
kyo	<input type="text"/>
all	<input type="text"/>
vtr	<input type="text"/>
ptb	<input type="text"/>
flu	<input type="text"/>
fyi	<input type="text"/>
zcn	<input type="text"/>
dqe	<input type="text"/>
neo	<input type="text"/>

How is an applicant assigned to you?

After the 2nd step is finished (determination of the wage), every employer is going to draw a random number between 1 and 16. This number determines in which order the employers are assigned to their applicants. This means, the drawn number also influences your probability of employing a certain applicant.

Assignment and employment are set up in the following way: The employer who has drawn random number 1 gets assigned the applicant he has ranked no.1. This means, he can employ the applicant he mostly preferred. This applicant will be paid the wage her employer has chosen before and the applicant can determine her effort with this information. This applicant is also, up from now, not available any longer on the labor market for the other 15 employers. That is why she is canceled from the individual ranking orders of the remaining 15 employers. There remain 15 available applicants on the labor market. Now it is the turn of the employer who has drawn random number 2. He will be assigned the applicant who is ranked best on his ranking order of the remaining 15 applicants. This could be the applicant who is ranked no.1. But in case this specific applicant is already employed by the employer who was allowed to choose first, the applicant ranked no.2 will be employed. Afterwards, it is the turn of the employer who has drawn random number 3.

This procedure continues until each employer is assigned an applicant. Every employer receives from the remaining applicants the corresponding applicant with the best rank. As you can see it is important to carefully think about the rank order - not only for the first rank but also for all other ranks up to rank 16.

2nd step: Employer's wage-decision

Within this step, you have to decide which amount of wage you are willing to pay for an applicant in case of an employment. Since you do not know yet which applicant is going to be employed, you have to choose a wage for each of the 16 applicants. Be aware that every pay off might be relevant. For this decision you once again use the data sheets of the 16 applicants. The wage you want to pay needs to be whole-numbered between 0 and 100, that is

$$0 \leq \text{wage} \leq 100.$$

The entry mask on the computer for your wage decision looks like this:

2. Schritt: Lohnzuordnungsentscheidung

Kurzlebenslauf	Ihr zugeordneter Rang	Ihre Lohnzahlung
msä	4	<input type="text"/>
dlp	15	<input type="text"/>
akr	3	<input type="text"/>
hpl	13	<input type="text"/>
hpn	8	<input type="text"/>
ogp	5	<input type="text"/>
joz	7	<input type="text"/>
uoy	1	<input type="text"/>
trr	16	<input type="text"/>
ail	9	<input type="text"/>
vfr	10	<input type="text"/>
gfa	2	<input type="text"/>
flu	12	<input type="text"/>
esp	6	<input type="text"/>
zcn	14	<input type="text"/>
gfu	11	<input type="text"/>

After having completed the 2nd step, you will draw the random number which determines the assignment of employers and applicants according to the personal ranking orders. After having drawn the number and having entered it into a separate computer, the code of the applicant who is assigned to you appears on your screen.

After you got to know with whom you will interact, we will ask you to fill out a questionnaire. Afterwards, today's session will be closed for you.

3rd step: Decision of the applicants

Within this step, applicants choose their effort they are willing to contribute. Applicants will therefore first be informed about the fact that they are engaged by an employer. They will also get to know their amount of wage they are going to receive from their employer. Applicants do neither get to know which amount of wage the employer was willing to pay for other applicants nor whether the employer preferred other applicants.

Applicants can choose any whole-numbered effort between 1 and 10, that is

$$1 \leq \text{effort} \leq 10.$$

The effort shall be conducive to the benefit of the employer. The effort causes costs for the applicants. These costs depend on the amount of effort. The higher the effort is, the higher are the costs.

This step takes place at a different point in time after this experiment, within a separate experimental session.

How are incomes calculated?**Income of employers: Your income**

Your income depends on the wage you are paying your employed applicant and the effort the applicant has contributed for this wage. Your income is determined in the following way:

$$\text{Your income} = 10 \times \text{effort of the applicant} - \text{wage}$$

As it can be seen in the above formula, the higher the effort the higher your income is and the lower the paid wage the higher your income is.

Income of the applicants: Income of the applicant you have engaged

The income of your applicant corresponds to her wage minus the costs of her effort, she has to bear. The income of your applicant is determined in the following way:

$$\text{Income of your applicant} = \text{Wage} - \text{effort costs}$$

The higher the applicant's chosen effort the higher her costs are. In the following chart, costs for all possible efforts are listed:

Effort	1	2	3	4	5	6	7	8	9	10
Costs	0	1	2	4	6	8	10	12	15	18

The higher the applicant's wage the higher her income is and the lower her effort the higher her income.

All employers' and applicants' income are calculated in the same way. This means, each employer can calculate the income of his applicant and each applicant can determine the income of her employer.

Bear in mind that employers and applicants can also achieve deficits. In case you achieve a deficit at the end of the experiment, you have to balance it out of your endowment (=10 EURO).