On Peer Effects: Behavioral Contagion of (Un)Ethical Behavior and the Role of Social Identity

Dimant, Eugen

University of Pennsylvania, Behavioral Ethics Lab, Harvard University, Edmond J. Safra Center for Ethics

28 December 2015

ON PEER EFFECTS: BEHAVIORAL CONTAGION OF (UN)ETHICAL BEHAVIOR AND THE ROLE OF SOCIAL IDENTITY

Eugen Dimant[‡]

University of Pennsylvania, Behavioral Ethics Lab

Harvard University, Edmond J. Safra Center for Ethics

Version: December 2015

ABSTRACT:

Social interactions and the resulting peer effects loom large in both economic and social contexts. This is particularly true for the spillover of (un)ethical behavior in explaining how behavior and norms spread across individual people, neighborhoods, or even cultures. Although we understand and observe the outcomes of such contagion effects, little is known about the drivers and the underlying mechanisms, especially with respect to the role of social identity with one's peers and the (un)ethicality of behavior one is exposed to. We use a variant of a give-or-take dictator game to shed light on these aspects in a con-trolled laboratory setting. Our experiment contributes to the existing literature in two ways: first, using a novel approach of inducing social identification with one's peers in the lab, our design allows us to analyze the spillover-effects of (un)ethical behavior under varied levels of social identification. Second, we study whether contagion of ethical behavior differs from contagion of unethical behavior. Our results suggest that a) unethical behavior is more contagious, and b) social identification with one's peers and not the (un)ethicality of observed behavior is the main driver of behavioral contagion. Our findings are particularly important from a policy perspective both in order to foster pro-social and mitigate deviant behavior.

KEYWORDS: Conformity, Behavioral Contagion, Peer effects, Social Identity, Unethical Behavior

JEL: D03; D73; D81

[‡] This is the full version of the paper with an extensive literature review published in my dissertation entitled *'Economics of Corruption and Crime: An Interdisciplinary Approach to Behavioral Ethics'* (2015). This work has greatly benefited from conversations with Edward Glaeser, Daniel Houser, Lawrence Katz, David Laibson, Ulrich Schmidt, and Wendelin Schnedler on the early version of the experimental design. I am particularly thankful for valuable input from Max Bazerman, Cristina Bicchieri, Gary Bolton, Elena Katok, Judd Kessler, and Robert Kurzban during my visiting research positions at the Harvard University, the University of Pennsylvania, and the University of Texas at Dallas, respectively. I also want to thank René Fahr, Uri Gneezy, Burkhard Hehenkamp, Rosemarie Nagel, Arno Riedl, and Tim Salmon as well as the participants at numerous workshops and conferences for interesting discussions and helpful suggestions. Financial support by the German Research Foundation (DFG) through the SFB 901 at the University of Paderborn is gratefully acknowledged.

Table of Contents

1. Introduction

Individuals do not co-exist in pure isolation but interact within social contexts, or as the eminent Elliot Aronson (2011) emphasizes it in his book's title: individuals are *social animals*. Despite the long tradition in anthropology and sociology, economists have been rather negligent of the relevance of norms, values, and social influence of peers on one's behavior for a long time. Fortunately, over the last two decades there has been a push in the economic discipline to expand our understanding of what comprises a more sophisticated individual decision maker by accounting for the individual's identity, morals, and other-regarding concerns (cf. Rabin (1993), Fehr & Schmidt (1999), Akerlof & Kranton (2000), Bolton & Ockenfels (2000), Charness & Rabin (2002), Bénabou & Tirole (2011)). These approaches have enriched our understanding by regarding social and economic decisions as a function of the respective social and economic environment and the relevance of one's peers' behavior.

Recently, both economists and psychologists have started engaging in a promising dialogue on behavioral ethics and the drivers of (un)ethical behavior by bringing together classical and behavioral approaches (for a recent review see Irlenbusch & Villeval (2015)). A standard economic argument is the assumption of fixed preferences, translating into one's conforming behavior being the result of social conventions or norms. Conversely, psychologists, sociologists, and recently some economists, among others, have challenged this fundamental assumption, suggesting that behavioral adaptation is the result of converging preferences and fluid tastes. Bernheim & Exley (2015) refer to the former as *belief mechanisms*, and to the latter as *preference mechanisms*. Along these lines, scholars in economics and psychology have attempted to shed light on the mechanism of peer effects using both lab and field experiments (for an overview see Houser et al. (2012) and Lahno & Serra-Garcia (2015)). Exemplarily, peers are found to significantly affect individual judgment (Asch, 1951) and risk taking behavior such as credit decisions (Banerjee, et al., 2013), stock market participation (Shiller, 1984), investment decisions (Scharfstein & Stein, 1990), and littering behavior (Cialdini, et al., 1990). Peer effects are also at play in education (Sacerdote, 2001) and productivity at work (Falk & Ichino (2006), Mas & Moretti (2009)), and in the form of neighborhood effects (Case & Katz (1991), Katz, Kling & Liebman (2001), Kling, Ludwig & Katz (2005), Kling, Liebman & Katz (2007), Chetty, Hendren & Katz (2015)).

Studying behavioral spillovers and the adaptation of observed peer behavior particularly in the domain of (un)ethical behavior (we will call it *behavioral contagion* or *behavioral adaptation*) looms large due to its economic and social significance. Much like emotional contagion, the underlying idea of behavioral contagion depicts a form of social influence leading to the emulation of behavior one gets exposed to (see Wheeler (1966) for a discussion and differentiation from other frequently used terms such as conformity and imitation). It comes natural to analyze situations in which individuals can simultaneously engage in either ethical or unethical behavior. However, most existing studies have rather focused on analyzing ethical and unethical behavior in isolation. Economic studies have highlighted peer-effects in pro-social behavior (Frey & Meier (2004), Gächter et al. (2013)), voluntary cooperation (Thöni & Gächter, 2015), as well as within the unethical domain, such as the use of performance enhancing drugs (Gould & Kaplan, 2011), and dishonesty (Innes & Mitra, 2013). In sum, there exists ample evidence that peer effects are a phenomenon across different contexts, social environments, and cultural groups and individuals adjust own behavior to resemble one's peers.

However, scholars across various fields are still disunited on whether the methods used qualify to observe clean peer effects, or whether for a large part our observations are an artifact of potential confounds. Exemplarily, a stream of literature points at methodological problems in soundly measuring such effects, especially outside the controlled laboratory environment (for a critical discussion see Manski (1993) & (2000) and Angrist (2014)). Our work adds to this literature by broadening our understanding of how peer effects and the resulting behavioral contagion play out within the spheres of both ethical and unethical behavior simultaneously. Using a novel design that allows us to study different types of behavior and behavioral spillovers simultaneously is one of the ways we add to the existing literature. Beyond that, we use a novel approach to induce varying levels of social identity in the lab to study behavioral contagion in different social settings. We will return to this shortly.

Our goal is to contribute to this debate by shedding light on two aspects: first, does social identification to one's peers facilitate the spread of (un)ethical behavior? Second, is the magnitude of behavioral contagion dependent on the (un)ethicality of the behavior and if so, which behavior is more contagious? By answering these questions, we contribute to the growing economic literature on the role of identity and social context on behavior in general (cf. Hoffman, McCabe & Smith (1996), Bohnet & Frey (1999), Akerlof & Cranton (2000), Charness & Gneezy (2008)) and spillovers resulting from peer effects in particular (i.e. see the previously

mentioned seminal Moving-to-Opportunity literature). However, in our study we analyze lower-bound peer effects because adaptive behavior remains unobservable by one's peers and thus carries no signaling value. Because such a setting sets us apart from what is typically meant by the term *conformity*, we use the more encompassing term *behavioral adaptation* (for a discussion of the mechanism and empirical literature on conformity see Bernheim (1994) and Bikhchandani, Hirshleifer & Welch (1998)).

Embedded in a controlled lab experiment, our approach to measuring peer effects is a variation and extension of a dictator game as introduced by List (2007) and used by Bardsley (2008). We capitalize on a one-shot dictator game in which participants are given the opportunity to give to or take money away from the charity before and after learning peer behavior. That is, the extend the one-shot setting by introducing a revision option to account for behavioral contagion (see Thöni & Gächter (2015) for a related approach). We deal with the noted reflection problem (Manski, 1993) by introducing a novel design, which we will discuss in more detail in the design section of our paper (see chapter 6). In short, our approach centers on two key design elements: first, only those who *actively observe* another participant's behavior can react and revise initial behavior. Second, behavior of those who *are observed* is held fixed and cannot be changed afterwards, which is public knowledge. Such an approach allows us not only to study general peer effects in an unbiased way, but also to shed light on the relevance of factors such as the social identity to one's peers and the (un)ethicality of observed behavior in driving behavioral changes. Such an experimental set-up allows us to contribute to existing research on peer effects in multiple ways and thus open a venue for future research.

In particular, we construct a list of personal statements taken from a major American dating website to categorize participants according to overlaps in preferences and interests (for a discussion, see Hitsch, Hortaçsu & Ariely (2010)). [2] This approach allows us to create a measure of social proximity and use this as an exogenously varying matching device to study peer effects in the lab and to combine the best of both worlds: for one, the controlled laboratory setting mitigates the previously discussed arising difficulties when studying peer effects in the field. For another, our novel measure of social proximity allows us to mimic social affection that is normally absent in the laboratory setting due to its intended nature of anonymity. Our

[2] We use these questions to generate a composite matching score with other participants and randomly vary the information set across treatments given to matched pairs of participants. The exact matching mechanism and the treatment variations will be explained in more detail in chapter 6.

results indicate the validity of this proximity measure as the observed peer effect (i.e. the magnitude of the behavioral reaction) is aligned with the degree of measured social proximity.

Our study extends the existing literature on peer effects in a number of ways. We do not only provide a clean approach to test *whether* peer effects exist in the context of both ethical and unethical behavior; we also deliver sound evidence on *how* such peer effects depend on the social identification with their peers. We achieve this by contributing from both the content and the methodological perspective.

From a content perspective, our work focuses on making a substantial contribution to better understand drivers of behavioral adaptation in the domain of both ethical and unethical behavior. In social dilemmas and cooperative settings, existing theoretical and experimental literature points out that giving and taking are indeed different concepts that trigger different behavior (cf. Cox et al. (2008), Gächter et al. (2015)). Essentially, we put these findings to a test in a setting of individual decision-making under peer influence in which no strategic interaction between individuals exists. That is, we try to answer whether the unethicality of observed behavior or social identification to one's peers is a stronger predictor of behavioral contagion. While different streams of research suggest the relevance of both channels (see discussion in the next chapters), we are the first, to the best of our knowledge, to evaluate both channels in terms of their impact on behavioral contagion. From a methodological perspective, our contribution in this paper is to introduce an approach that does not only allow us to directly mimic social proximity in the lab, but also to vary its degree. If successful, such an approach opens the venue for a broad range of more refined future research examining the role of social proximity in cooperation, reciprocity, and punishment behavior.

So far, the economic literature in particular has been fairly silent on answering three naturally arising questions with respect to peer effects that we will attempt to answer in this paper.

Question 1: *Is there a systematic difference of behavioral contagion with respect to unethical behavior (e.g. taking away) as compared to ethical behavior (e.g. donating)?*

It is reasonable to assume that behavioral contagion is a function of the observed behavior's nature. One obvious reason is the costs involved: good behavior implies bearing costs in order to improve the well-being of others, while bad behavior often implies improving one's own

well-being in one way or the other at the expense of a third party. A number of seminal empirical contributions examining neighborhood effects indicate that the adaptation of behavior is asymmetric and depends on whether one has been exposed to good or bad influences and to what extent (cf. Kling, Ludwig & Katz (2005)). While existing studies have typically resorted to explaining behavioral contagion in one direction, that is either ethical or unethical, our experimental design allows us to compare the contagion of both behavioral domains directly. To our knowledge, this is the first controlled approach to directly and simultaneously compare peer effects and behavioral spillovers across these two domains.

Question 2: *To what extent is behavioral contagion in either direction (i.e. of moral and immoral behavior) mediated by the social proximity to the peers?*

Following the existing literature on social identity, it is reasonable to assume that observing behavior of people who are socially closer or similar depicts a more salient signal in terms of what is socially accepted or an existing norm. However, the exact interaction between social proximity and the contagion of (un)ethical behavior remains unclear. In addition, good and bad behavior differs in terms of the information set available to the individual. While good behavior might entail some ambiguity with respect to what is 'appropriate' within a given context, bad behavior might be less ambiguous: the nature of bad or unethical behavior implies the overstepping of (social) boundaries or infringing laws. Consequently, the wiggle-room for self-justification is narrower in the latter case.

Question 3: *What is the stronger driver of behavioral contagion – social proximity or the (un)ethicality of observed behavior?*

Two quite distinct streams of literature exist that examine the role of either social proximity or exposed behavior within the framework of peer effects: one on the relevance of social identification, and one on behavioral observation. In this research, we attempt to unify both lines of research by using an experimental design that encompasses both aspects. We will return to this aspect in more detail in the next chapters.

In anticipating the results, we find that the magnitude of spillover-effects is a function of social identification that is asymmetrically biased towards the contagion of unethical behavior. Overall, our results suggest that within a given peer context it is more likely to observe behavioral contagion in the form of unethical than ethical behavior. Across different specifications,

we find that social proximity to the peers is more relevant to the crowding-out than to the crowding-in of ethical behavior, while the observation of (un)ethical behavior alone is insufficient to trigger any particular behavioral change. Thus, the mere observation of behavior alone is insufficient for the existence of peer effects in the (un)ethical sphere, but is rather contingent on the social identification to one's peers. The interaction between social identification and type of observed behavior adds to the understanding of peer effects and yields relevant policy recommendations.

In summary, it can be stated that our experimental work is along the lines and an extension of the seminal Moving-to-Opportunity (MTO) field studies that examine neighborhood effects and assimilation of behavior (see discussion in chapter 3.3). To the best of our knowledge, our paper is the first experimental examination of behavioral spillovers as a function of varied levels of social identity and the (un)ethicality of observed behavior, which in combination with the methodological novelty depicts our principal contribution.

The paper is structured as follows: Section 2 provides a more detailed thematic background and the course of our investigation, whereas we deal with the conceptual framework of behavioral contagion and its existing relevant literature in more detail in Section 3. In Section 4, we will discuss the drivers of behavioral adaptation and some of the more relevant concepts from the fields of economics and (social) psychology. We will introduce a simple theoretical model in Section 5 before turning to our experimental analysis in Section 6. We close with a concise discussion on potential policy recommendations in Section 7 and a conclusion and an outlook in Section 8.

2. Background and Course of Investigation

A basic principle of classical economic theory suggests that individuals form rational expectations based on available information and act on them accordingly. However, even the great John Maynard Keynes (1936) expressed his concern about the rationality of individuals to realize efficient investment decisions in the long run already 80 years ago. Instead, Keynes expected individuals to follow the herd, thus stressing the importance of peers for many economic decisions. Since then, a contrasting strain of literature emerged that accounts for the relevance of behavioral traits on the individual decision-making process. Understanding the underlying mechanism of peer effects is key to comprehending its impact on economic decisions and outcomes. For the bigger part, existing research on peer effects mainly resorts to

field experiments or purely observational studies that are generally inferior to controlled lab experiments in terms of, among others, a clean identification of the relevant channels, endogeneity, and reflection problems (see Manski (2000), Falk & Fischbacher (2002)). Only recently, there has been a push to study peer effects in the lab, allowing us to gain a deeper and often a more reliable understanding of the underlying mechanism Angrist (2014).

Social interactions in general and the potentially resulting peer effects in particular play an instrumental role from both the societal and economic perspective. Existing literature indicates that standard economic forces alone cannot encompass many of the outcomes that we observe in real life. Examples are, among others, the escalation of crime rates or the massive surge in female labor participation rates in World War II (cf. Mulligan (1998), Levitt (1999). See also recent findings on paternity leave by Dahl, Løken & Mogstad (2014)). Instead, social interactions are found to offer explanations helpful to understanding the causes of rapid shifts in economic fundamentals. Such ripple effects are likely the result of social interactions, thus raising the awareness about the importance of understanding the underlying mechanism of peer effects (Glaeser & Scheinkman, 2004).

Although explicit research on peer effects and the resulting behavioral spillovers (in the literature sometimes referred to as behavioral or social contagion) has its origins in the late 19th century, the underlying concept has been observed long before.[3] Reportedly, an abstruse-seeming stream of suicides happened after reading Goethe's *The Sorrows of the Young Werther* two hundred years ago. "My friends […] thought that they must transform poetry into reality, imitate a novel like this in real life and, in any case, shoot themselves; and what occurred at first among a few took place later among the general public […]" (Goethe, quoted in Rose (1929, p. 29)). The widespread imitation of this behavior gave rise to fear among the population and governments, ultimately leading to a ban of the book in Italy, Leipzig, and Copenhagen (Phillips, 1974). The outbreak of the Tanganyika laughter epidemic of 1962 in Uganda is another infamous example of behavioral contagion. There, a mass hysteria infected

[3] Research on behavioral contagion is fragmented and different disciplines have introduced own notions and definitions referring to the same or closely related concept. Existing research interchangeably uses different terms to describe such situations, among others: conformity, behavioral contagion, imitation, or behavioral adaptation. Due to its more generic nature and the context of our research, we will mainly resort to the term behavioral contagion or behavioral adaptation. For the sake of comprehensibility, we will abstain from clearly defining and delimiting those concepts for now. We will return to this point in chapter 2.

almost 100 pupils with contagious laughter, forcing several schools to close down for days (Rankin & Philip, 1963).

Initially, the concept of social contagion has been introduced in the form of a social phenomenon – as opposed to a biological one – explaining why and how certain forms of behavior soak through society (for early work see Baldwin (1894), Tarde (1903)). Since the 1950s, empirical research on this topic has been on the rise with evidence suggesting that the mere exposure to and contact with individuals or culture is sufficient to trigger behavioral contagion. Conditional on a sufficiently salient trigger, behavioral contagion leads to behavioral adaptation towards observed behavior. In this paper, we aim at expanding the existing knowledge on the drivers of saliency, in particular with a focus on both the ethicality of the observed behavior and the degree of social distance or proximity with the observed individual. A long tradition in social science highlights the importance of social identity in understanding individual behavior within the framework of social interactions (cf. Bogardus (1928)). A salient state of social identity is found to trigger favoritism towards those of stronger social kinship.

The term social identity is eclectic and several of its facets have been studied in existing economic research. While the term encompasses a broad range of conceptual elements, from shared preferences and experiences to shared cultural and religious beliefs, in this paper we follow the primal understanding of this term as introduced by Tajfel and Turner (1979). More precise, we refer to the existence of social identity if a person derives self-esteem from belonging to the peer group and has a preference for exhibiting similar behavioral patterns (for a similar approach, see Chen & Li (2009)). Using this definition in combination with a simple but richer implementation of social identity, as applied in our experiment via the observation of preference similarity, is conducive to both deriving lower bound results for the role of social identity in facilitating (un)ethical behavior and easier reproducibility of our results. This line of research is important from a policy perspective in generating effective measures to trigger both more pro-social and less anti-social behavior, consequently reducing the otherwise resulting economic and social inefficiencies.

With notable exceptions, existing research has been struggling to overcome a number of challenges to study clean peer effects in the lab, especially in contexts where social identity plays a mediating role. Among these, inducing or at least proxying the natural occurring variation of

social identity has proven to be difficult. Following the tradition of the minimal-group paradigm (Tajfel & Turner, 1986), psychological research has introduced a number of ways to proxy social identity in the lab, such as having participants interact with participants that were assigned the same color avatar (Tajfel, 1982), a similar name (Pelham, et al., 2005), same birthday (Cialdini & DeNichols, 1989), or imagined closeness (Gunia, et al., 2009). So far, it has proven to be challenging to study peer effects under controlled settings in general, let alone within a more sophisticated social environment such as varied levels of social identity. Existing economic research has resorted to using a dual approach, studying peer effects in the lab and in the field, with both approaches having their limitations. Applying a novel methodological approach to induce varying levels of social identity in the lab allows us to combine the best features of both worlds to study peer effects beyond what had been possible so far.

To our knowledge, the first economic contribution examining this question in a controlled setting by varying social identity among individuals is the study by Bohnet & Frey (1999).[4] They used a rough proxy for social distance by varying the degrees of identification and found evidence that social distance is decisive in predicting the extent of other regarding behavior in a dictator game setting.[5] Other examples of economic approaches that bridge anonymity and induce social identity, among others, vary the wording of the experimental instructions (Hoffman, et al., 1996), use face-to-face interaction (Bohnet & Frey, 1999) or show pictures to one party only (Eckel & Petrie, 2011), reveal names (Charness & Gneezy, 2008), reveal preferences such as those for paintings (Chen & Li, 2009), or recruit friends and family members (Brandts & Solà, 2010). Such studies typically yield the robust finding that stronger social identity triggers favoritism. Two natural problems arise with the concepts used in economics so far. Firstly, using face-to-face communication or allowing participants to interact with friends or family members introduces serious biases and crowds-out the revelation of true preferences (Roth, 1995). Secondly, and more relevant to the point of our experiment, the degree

[4] Although motivated by a previous study of Hoffman, McCabe & Smith (1996), the work of Bohnet & Frey (1999) is the first to directly vary (a proxy for) social distance among peers. Instead, Hoffman, McCabe & Smith (1996) varied the language used in the distributed instructions, arguing that "subjects bring their ongoing repeated game experience and reputations from the world into the laboratory, and […] dictator instructions […] may imply that the objective is to share the money with someone, who, though anonymous, is socially relatively near to the decision maker" (Hoffman, et al., 1996, p. 655).

[5] Research usually refers to this concept as social identity, which encompasses both social distance and its inverse, social proximity. Throughout this paper, we will mainly refer to social proximity. Another approach to mimic social proximity used in economic and psychological experiments has been to ask participants to bring along their friends or relatives and study their interaction in a controlled environment. Among other things, we will discuss potential drawbacks of these and related approaches in chapter 5 in more detail.

of social identification measured in the lab by these concepts can hardly be varied and is rather binary. Our approach of using dating-website questions as a matching device allows us to induce and exogenously vary different levels of social proximity in the lab in order to study their role in the spillover of (un)ethical behavior.

It is this paper's aim to shed light on three questions stated above and to contribute to a better understanding of the general mechanism of peer-effects. For this reason, we propose a novel approach to proxy different levels of social proximity among peers in a laboratory setting. Such an approach allows us to exogenously vary social characteristics and study their role in behavioral contagion. We mimic social proximity by the use of questions taken from a major American dating website to capture individual preferences and interests and use the matching scores of overlapping answers among lab participants as an exogenous matching device across treatments. This allows us to study decision-making going beyond simple ingroup - outgroup comparisons. Rather, our approach provides us with an extensive array of possibilities to match participants according to their shared similarities. To the best of our knowledge, we are the first to use such an approach. Thus, we not only complement existing field studies, but also broaden the scope and utilization of lab experiments in explaining behavior and behavioral changes in peer settings, especially within the unethical domain.

For this purpose, we extend the currently existing approaches by a social component that sufficiently considers the relevance of social distance and proximity to one's peers in affecting behavioral decisions. This attempt will be at the heart of this paper, leading to a proposed theoretical extension of Akerlof's (1997) and Glaeser & Scheinkman's (2004) seminal work on social interaction, social distance, and conformity, where the extent of social distance is a function of geographic location. We, however, emphasize the role of social distance as a function of an actual overlap in personality based on e.g. personality traits or interests.

Another substantial contribution of our research is the direct comparison of behavioral contagion of ethical and unethical behavior. In general, existing research has focused on shedding light separately on behavioral spillovers of either ethical behavior (cf. Thöni & Gächter (2015)) or unethical behavior (cf. Gino et al. (2009)). Considering the differing settings and games used in existing experiments to study behavioral contagion, current research does not help to understand whether and to which extent behavioral contagion in either direction differs from each other. Instead, experiments that add to the understanding of the potentially different

mechanisms should place participants in a uniform environment and allow for the spillover of ethical and unethical behavior.

Our experimental set-up allows us to study these questions. Participants play a one-shot dictator game in which they decide how much money to donate to or take away from a charity, which resembles a variation of a dictator game implemented by List (2007) and Bardsley (2008)). Here, the (un)ethicality of (taking away) donating money to the charity is stressed by explaining the consequences of their behavior clearly to the participants. That is, all the money that is (taken away) donated to the charity will (not) be given forward to the charity, thus individual behavior will (harm) benefit the charity. Hence, participants face a riskless but in terms of its (un)ethicality precisely defined situation in which they have to decide whether or not to personally benefit at the expense of a charity of their choice. After reaching their initial decision, participants are given the opportunity to learn about other participants' initial decisions followed by the option to revise their own initial decision.

Several economic and psychological theories are able to explain behavioral contagion even under full anonymity and without observability of one's own initial and potential revision decision, as implemented in our experiment. Among these are concepts relating to *social decisions and social distance* (Akerlof (1997), Glaeser & Scheinkman (2004)), *imitation of preferences* (Sliwka, 2007), *social learning* (Bandura, 1971), *norms* (Cialdini et al. (1990) and Bicchieri (2006)), *self-expansion* (Aron & Aron, 1986) or even guilt (Kandel & Lazear, 1992). Many of these concepts are not strictly distinct, in both their assumptions and predictions. Thus, in this paper we will not attempt to resolve which approach explains behavioral contagion best but rather focus on shedding light on the drivers of behavioral contagion and its interrelation with the social identity dimension and the extent of (un)ethicality of observed behavior.

3. The Conceptual Framework of Behavioral Adaptation

3.1 The Mechanism of Social Interaction and Challenges of Measuring its Effects

A growing body of literature suggests that social interactions are principal not only to humans, e.g. in making social or economic decisions, but also to animals, e.g. in finding the right strategy or place to maximize one's hunting success (Laland, 2002). Evidently, social interactions trigger different reactions and outcomes, which are referred to as social effects. Different

streams of literature refer to such interactions in different ways: bandwagons, conformity, epidemics, herd behavior, imitation, neighborhood effects, peer influences, social learning, or social norms (see Hyman (1942), Merton (1957), Granovetter (1979), Jones (1984), Manski (2000)). In addition, Pingle & Day (1996) subsume these and other types of behavior, including following an authority, habit, thoughtless impulse, and hunch, as *economizing behavior*. Some call it simply peer effects.[6] Although these notions refer to different mechanisms, their outcome is often (but not always) similar; that is, the individual's adaptation to observed behavior. However, some mechanisms encompass stronger interaction with the peers than others do (e.g. the peer's ability to observe my reaction and exact change in behavior) and involve more or less deliberation.[7]

As we will discuss in chapter 3.3 in more detail, existing research indicates that the existence and persistence of social effects are context specific and may even lead to, among others, higher consumption of alcohol and drugs, cheating, and smoking. The same is also found to be true for pro-social and cooperative behavior. Thus, before turning to the specific effects of social interaction, we shall make the effort to understand the underlying mechanisms first.

The concept of social interaction lies at the heart of social psychology and sociology, for example in order to explain the formation of tastes (Weber, 1978). The importance of social interaction in explaining social phenomena has a long tradition and is often ascribed to Sutherland's Differential Association Theory (Sutherland, 1939). Akerlof (1997) stresses the view that the theory of social interaction is key to understanding why individuals do not succumb to isolated and purely self-maximizing decision-making. Instead, this concept gives rise to conceive an individual as someone who constantly interrelates with the underlying social environment and produces and deals with the resulting externalities. A principle consequence of

[6] Often, the term "peer effects" is used to refer to all of these mechanisms without specifying the exact channel through which behavioral adaptation arises. Although it is important to shed light on the different channels through which social interaction potentially transitions into behavioral adaptation, the focus of this study lies in understanding the role of exogenous factors such as social proximity that have the potential to influence the intensity of behavioral adaptation. We shall not attempt to settle the argument of which approach explains the mechanism of our experiment best. Instead, we will discuss some of the more prominent concepts and mechanisms in the fields of economics and social psychology that resemble the mechanism of behavioral adaptation the way it is implemented in our experiment in chapter 4. For our purpose, however, we will assume peer effects to be in play whenever an individual $\boldsymbol{i}'s$ behavior changes after having been exposed to behavior of individual $\boldsymbol{j}$, irrespective of the direction of the behavioral change. In turn, however, we assume behavioral contagion to be in place whenever individual $\boldsymbol{i}'s$ behavior changes in the direction of individual $\boldsymbol{j}$ behavior.

[7] It is worth noting that it is far beyond this paper's scope to shed light on all of these concepts. Instead, we will pick out and discuss those concepts of which we believe are of bigger importance to what we analyze within our experiment. We will return to this point in chapter 4.

extending the rational model based on Becker's early work by such a social dimension is that the type of the resulting individual is more sophisticated and resembles more closely to the intuition of sociologists than the classical economists.

Manski (1993) distinguishes between two types of social interactions.[8] One is endogenous interaction in the form of, for example, information exchange among criminals or social norms. The behavior of the relevant peer group mediates the likelihood that the individual will engage in the same kind of behavior. In addition, exogenous interaction emphasizes that the propensity of an individual to behave in a certain way is also mediated by exogenous characteristics of the group such as their attitude toward crime or social and economic status.[9]

In social interactions, externalities abound. The key mechanism of social interaction implies that one's personal net benefit is a function of the behavior exhibited by one's relevant social group or contact person (Glaeser & Scheinkman, 2004).[10] Inherent to such interactions are strategic complementarities in the form of circular cascades where "even if changes in fundamentals create only a small change in the level of activity for each individual, each individual's small change will then raise the benefits for everyone else pursuing the activity" (Glaeser & Scheinkman, 2004, p. 84). In principle, small changes in fundamentals may cause large shifts in outcomes, which are sometimes referred to as the *butterfly effect*, a term hailing from the chaos theory (Lorenz, 1963).

Social interactions are also highly relevant in understanding the spread of criminal behavior. Social interaction plays a decisive role in the formation of gangs and the recruitment of young criminals (see Reiss (1988) and Jankowski (1991)). This is particularly true for the criminals'

[8] Glaeser & Scheinkman (2004) provide a more distinct categorization of the mechanisms that generate social interaction: physical, learning, stigma, and taste-related interactions. For the purpose of this experiment, we will extend this categorization in order to better capture the mechanism of behavioral changes we are interested in. We will discuss these points in more detail in chapter 3.

[9] While these two aspects represent interactions that are shaped by the underlying social environment, Manski (1993) also introduces correlated effects that explain similar behavior as the result of facing similar institutional environments (see also Manski (2000)). This third aspect is not considered any further since correlated effects are not social effects and are thus neither created by social interactions nor create social multipliers (see Glaeser, Sacerdote & Scheinkman (1996), (2003)). Using a novel approach in our experiment, a clean variation of both endogenous and exogenous interaction allows us to draw causal inferences that are more precise than what has previously been possible. We will return to this important point in our design section of chapter 6.

[10] The term 'relevant social group or contact person' is vague in existing research. A social group or contact person is self-reported and defined from one's individual point of view and refers to one or more individuals whose behavior either has a direct or indirect impact on one's well-being, e.g. through resource externalities or other-regarding preferences. While it is important to understand the different channels through which behavioral spillovers occur, the focus of our study is to shed light on the drivers instead. See chapter 4 for a discussion.

decision to engage in illicit behavior jointly (Reiss (1980)). "Social interactions seem to create a sense of invulnerability and a willingness to violate social norms and take risks, as long as one is in the company of like-minded individuals" (Glaeser, et al., 1996, p. 511). Social interaction also strongly affects stigmatization, which is important from an evolutionary perspective. As the number of criminals rise, illicit behavior becomes more common and thus potentially more accepted; and as criminality converges towards 'normality', the (social) rents of illicit behavior increase due to increased attractiveness, social acceptability and a crowding-out of legal activities where earnings from legal activities are stolen by criminals (see Rasmussen (1996) and Murphy, Shleifer & Vishny (1993)).

Researchers face substantial difficulties measuring social interactions and its resulting effects in a clean way. In a real-world instance, assigning changes in individual conformity to distinct mechanisms are subject to identification problems (Angrist, 2014). As argued before, this problem arises because individual behavior is affected by both endogenous (e.g. the group's behavior) and exogenous effects (e.g. group characteristics) and, in addition, uniform behavior can be the result of similar unobserved characteristics (Manski, 1993). Previous empirical research involved regressing a person's actions on the action of his peers. However, Manski (1993) points at three fundamental problems concerning this methodology: first, drawing causal inference is difficult when endogeneity is a problem, in particular when the individual's and its peer's behavior is interactive and influences each other circularly. Second, omitted variables increase the likelihood of spurious correlations between actions. Third, in reality, sorting and self-selection into particular neighborhoods renders it difficult to understand what actually drives behavior. Arguably, empirical research in particular faces these challenges because one only observes the behavior of individuals who self-selected themselves into, for example, moving to a better neighborhood, but not of those who decided to turn down the opportunity (Glaeser & Scheinkman, 2004).[11]

Of particular interest to our research are social interactions leading to the spread of unethical behavior. Existing research points to a strong presence of positive covariates across individuals' decision to engage in criminal behavior. In particular, Glaeser, Sacerdote & Scheinkman

[11] Prominent examples and forceful ways of addressing these issues include, among others, the work of Case & Katz (1991), Katz, Kling & Liebman (2001), Angrist & Lang (2004), Kling, Ludwig & Katz (2005), Kling, Liebman & Katz (2007), Ludwig & Kling (2007), Damm & Dustmann (2014), and Chetty, Hendren & Katz (2015). We will discuss these and other empirical contributions examining peer effects and behavioral adaptation in the field in more detail in chapter 3.3.

(1996) state that only covariance across criminal decisions of individuals explains existing variance in crime rates, which is far beyond any theoretical prediction of crime rates.

In order to mitigate any term-related confusion on the side of the interested reader, we are in need of a term that allows us to capture a particular type of behavior that resembles a subset of social interaction and its resulting social effects. More specifically, we are not interested in the drivers of all kinds of behavioral changes resulting from observation, but only in behavioral changes that lead to a convergence of behavior. We are thus proposing the impartial term *behavioral adaptation* to capture such behavior.[12]

The term behavioral adaptation refers only to a subset of social interaction because social interaction may lead to all kinds of social effects where the resulting behavior may or may not converge towards what has been observed. In its basic form, the existence of social effects does not tell us much about the specific behavioral reaction, if any, that follows from this exposure. In principle, the result of social interaction could lead to either behavioral alienation or adaptation. For the purpose of this study, we will focus on the latter. In turn, behavioral adaptation refers to those situations only in which the resulting behavior is coherent to the behavior one has been exposed to. Our study sheds light on the key drivers triggering an individual's response to become more like others, in one way or another.[13]

Existing macro- and micro-level data inadequately catches the underlying mechanisms and the causal relationships relevant to our project. That is, answering the questions of how behavioral adaptation varies with different levels of social proximity and whether this interplay is different for adaptation towards ethical versus unethical behavior. As has been argued by Angrist (2014), with rare exceptions like Kling et al.'s (2007) Moving-to-Opportunity research,

[12] Behavioral adaptation has also been studied in the field of evolutionary game theory as well as in the theory of learning in games (see Selten (1978), Roth & Erev (1995), Schlag (1998), and Apesteguia, Huck & Oechssler (2007), to only name a few). However, it is beyond the scope of this paper to touch upon all streams of literature that have shed light on the general process of behavioral adaptation. Instead, we focus on the studies most relevant to our experiment and extend our apologies to colleagues whose research remains unnamed in this paper.

[13] Studying the drivers of behavioral alienation is a potential venue for future research, as this line of research has yet to catch attention from economists. Akerlof (1997) refers to this behavior as the result of status-seeking efforts. Beyond this, however, behavioral alienation can be the result of e.g. one's desire to not be identified with a particular social group.

most field studies on peer-effects suffer in one way or the other from endogeneity, self-selection and confounds resulting from confusing the relevant subjects with the peers.[14] This renders it difficult to disentangle causation from simple correlation. For these reasons, Manski (2000) and Angrist (2014) emphasize the importance of controlled laboratory experiments that are able to reduce potential confounds to a much stronger degree, which is what we attempt to do here.

3.2 Why Understanding Behavioral Adaptation Matters

Social interaction and more so the resulting social effects tell us a lot about the underlying mechanisms affecting social and economic outcomes. As argued by Akerlof (1997, p. 1006), "social interaction theory explains why social decisions – such as the demand for education, the practice of discrimination, the decision to marry, divorce, and bear children, and the decision whether or not to commit crimes – are not simple choices based primarily on individual considerations." Along these lines, behavioral changes are subject to spillover-effects induced by observing a particular kind of behavior.

Arguably, economic literature neglected the relevance of spillover-effects on individual behavior for a long time and it is only recently that economists have begun incorporating motives beyond the neoclassical economic theory's assumption of own-payoff maximizing egoists (Akerlof, 1997).[15] It seems reasonable to assume that individual behavior is impacted by other people's behavior, their preferences, their sentiments and the like. As prominently stated by Fehr & Fischbacher (2005, p. 167), "if people believe that cheating on taxes, corruption and abuses of the welfare state are wide-spread, they themselves are more likely to cheat on taxes, take bribes or abuse welfare state institutions." Consequently, social interaction shapes one's own understanding of the world and behavior surrounding us.

Social interactions and its resulting effects are key in driving behavioral changes. In reality, we constantly make use of learning from and adapting to observed behavior. The rating systems of, among others, Amazon, eBay, and IMDB use the principle of social learning by having introduced a publicly accessible valuation system to spread the word of good and bad products and services. Instead of resorting to time consuming and potentially harmful trial-and-error

[14] For a rich and critical discussion on challenges relating to identifying and measuring social interactions, and avenues for future research, see Blume, Brock, Durlauf & Ioannides (2010).

[15] For early seminal contributions on the economics of altruism and egoism, see Becker (1976) and Becker (1981).

behavior, learning from others' experience is a survival strategy that facilitates (social) Darwinism and the stability of cultural cohesion (Laland, 2002). Evidently, social interactions lead to social effects that are likely to trigger a type kind of behavior that we are particularly interested in: behavioral adaptation.

Social effects are expected to be in place whenever events on an aggregate level interact with events on the individual level (Manski, 2000). Since the seminal contribution of Shelling (1973), a great amount of theoretical and empirical work highlighted the relevance of social effects in various contexts (cf. Evans, Oates & Schwab (1992), Glaeser, Sacerdote & Scheinkman (1996), Arcidiacono & Nicolson (2005), and Mas & Moretti (2009)). Such effects were also found to affect the individual's inclination to reciprocate behavior positively, consequently suggesting that social effects are of concern in trust relations (cf. Mittone & Ploner (2011)).

In the fields of psychology, economics and sociology there exists a long tradition emphasizing the impact of peers on individual behavior. Research efforts gave rise to a more extensive investigation of peer effects on behavior such as group norms (Sherif, 1936), bandwagon effects (Asch, 1951), conformity and social influence (Kelman, 1958), obedience to authority (Milgram, 1974), social dilemmas (Dawes, 1980), social norms (Elster, 1989), social networks (Wasserman & Faust, 1994), tax morale (Frey, 1997), moral identity (Akerlof & Kranton, 2000) and social capital (Putnam, 2000).

In sum, understanding why and how behavioral adaptation matters and how this is shaped by social interactions and its resulting social effects not only helps to understand the world around us but also facilitates the inception of policy measures that are more promising and target-aimed. Glaeser & Scheinkman (2004, p. 90) put it this way: "…if one person's level of education increases his neighbor's education through dissemination of learning, then it makes sense to subsidize education. There is a socially desirable spillover that should be subsidized. However, different policy implications appear if one person's level of education increases his neighbor's education for signaling reasons – namely, as one person gets more education the other person must also get more education or be thought inferior. In that case, there is a socially undesirable spillover that should not be subsidized."

3.3 What We Know: On Peers, Behavioral Adaptation, and Neighborhood Effects

Because the ability to induce salient social identity in an artificial lab setting is instrumental to the study of the posed interaction between social proximity and behavioral contagion, existing research has focused on field rather than lab experiments. However, Angrist (2014) argues that many field studies on peer-effects potentially suffer from endogeneity, self-selection, and confounds resulting from confusing the relevant subjects with their peers. Correlations might arise without any causation simply indicating a spurious relationship. Angrist (2014) points out that these challenges make it extremely difficult to disentangle correlation from causation and he thus calls for a controlled approach in the lab where relevant characteristics can be varied exogenously. In addition to problems arising from correlated unobservables and endogenous group membership, Manski (1993) also prominently coined the term "reflection problem", which results from the challenge of clearly disentangling the mutual influence peers exhibit on each other's behavior. Such a cyclical relationship between observed and actual behavior poses a huge challenge for studying peer effects in the field (Manski, 2000). Beyond identification issues, it has proven difficult to define appropriate peer groups and link them reliably to one another in the field.

This chapter is devoted to provide a state-of-the-art overview of existing research that has convincingly provided results on peer effects. We approach the discussion of existing research on behavioral adaptation by subdividing the literature based on its methodological approach, that is: evidence from the field or from the lab.[16] Although behavioral adaptation can be the result of various mechanisms, the economic literature (and especially its experimental subset) has put emphasis on studying peer effects, which is accepted as a term more broadly including the different mechanisms that we discussed previously. While this discussion is by far not exhaustive, we will concentrate mainly on experimental studies focusing on the relevance of peer effects in driving one's behavior that are most relevant to our study.[17]

Evidence from the Field

[16] Beyond question, previous research on peer effects not using experiments but rather observational data has been utterly important to today's research. However, for reasons discussed by Manski (1993) mentioned earlier, our literature discussion focuses on controlled field or lab experiments. For a discussion of empirical studies on peer effects, see Angrist (2014).

[17] In chapter 3, we will introduce and discuss different concepts from the fields of economics and (social) psychology that offer more distinct explanations to how and why (or why not) behavioral adaptation is driven beyond the encompassing term peer effects.

A comprehensive line of research suggests that peers decisively affect individual behavior. Several field experiments have investigated the change of individual contribution levels in response to the observation of other people's contribution decisions (cf. Frey & Meier (2004), Laundry et al. (2006), Croson & Shang (2008), Shang & Croson (2008)). However, Zafar (2011) argues that those results can be explained by at least two mechanisms working simultaneously: the individual's drive to conform to an underlying social norm or as a response to updated beliefs concerning the charity's quality.

For this reason, existing literature attempts to tackle the topic of behavioral spillovers from various angles to better understand the channels at work. For example, Ploner (2013) investigates whether peer's behavior affects intertemporal consumption choices at a university's cafeteria. He finds positive evidence for the existence of peer effects on individual decision-making. Mas & Moretti (2009) argue that peers substantially affect a worker's productivity levels positively (see also Azmet & Ichiberri (2010)). Bandiera & Rasul (2006) find that the farmers in Northern Mozambique condition their decision to adopt a new crop on the choices of their family and friends. Interestingly, they find an inverse U-shaped relationship suggesting that the observed social effects are positive if their social network contains few adopters and negative if a certain threshold is overstepped. The study of Sacerdote (2001) highlights that among college roommates, peers have an impact on grade point averages and the willingness to join fraternities. Ichino & Maggi (2000) find empirical evidence for shirking behavior within organizations, in particular for the case of a large Italian bank. They find a close relationship between an individual's absenteeism with the rate of absenteeism of co-workers. Cialdini et al. (1990) and Mas & Moretti (2009) show that the observation of another person's behavior leads to less littering in public places and higher productivity.

Along the lines of a more delinquent context, Wilson & Kelling (1982) have outlined the interdependence of disorder and criminality within a society, introducing the terminology of the 'Broken Windows Theorem'. Here, a broken window can function as a signal transmitting the understanding that social norms exist and tolerate fraudulent behavior. Social preferences and contextual information decisively affect one's understanding about what is seemingly appropriate in a given social context, thus shedding light on individual behavior from a comprehensive perspective (see Beckenkamp et al. (2014) for an experimental analysis). Another example for peer effects in the unethical domain is discussed by Gould & Kaplan (2011). Here,

the authors examined peer effects for the use of performance enhancing drugs for baseball players.

An extensive line of controlled experiment-in-the-field research focusing on neighborhood effects has been triggered by the seminal papers of Case & Katz (1991) in which they found evidence for criminal behavioral contagion both within families and neighborhoods in the Boston area. Ever since, multiple research projects have examined the short- and long-run effects of the Moving-to-Opportunity (MTO) project in which families are eligible to participate in a lottery for vouchers that would potentially help them to move to a better and safer neighborhood. Since 1994, this project has been in place in five cities: Baltimore, Boston, Chicago, Los Angeles, and New York City. Along these lines, Katz, Kling & Liebman (2001) examined the short-run effects of the MTO project on the well-being of the families who were offered a voucher. Their findings indicate a substantial improvement of well-being along different dimensions, including increased safety, and improved health conditions both mentally and physically. Surprisingly, especially young men were susceptible to the neighborhood change, while the young women's disobedience remained invariant.

With the ability to capitalize on a more extensive continuity of data, the studies by Kling, Ludwig & Katz (2005) and Kling, Liebman & Katz (2007) support previous results of an asymmetric assimilation process across gender. For criminal behavior, Kling, Ludwig & Katz (2005) find a strong gender effect. In terms of reduced arrests for violent crimes, men react positively to improved living conditions in the short-run. In the long-run, however, these effects vanish. Opposite to what can be expected from moving to a better neighborhood, males' general problem behavior and property crime arrest soar irrespectively. Conversely, females' criminal behavior decreases. The findings of Kling, Liebman & Katz (2007) indicate that although neither adult's economic self-sufficiency nor physical health conditions benefited from the MTO program, the mental health improvements for both adults and the female youth were substantial. What is more, the beneficial effects on education, risky behavior, and physical health that were found on the side of females were fully offset by the negative effects on the side of males, thus yielding limited contentment of the MTO initiative based on its overall impact. In a very recent study on the MTO program, Chetty, Hendren & Katz (2015) find strong evidence for positive income effects for children who were young (age 13 or younger) when their parents moved. Surprisingly, the same effects are either non-existent or even negative in the long-run for children who were older than 13 when they moved. The findings indicate that the

marginal gains from the MTO program decrease with the children's age, possibly due to disruption effects and social alienation. Although such findings are in line with literature suggesting that the duration of exposure of children to better environments is predictive of the treatment effect's magnitude, the results dampen the overall expedience of the MTO program (for a discussion see Chetty, Hendren & Katz (2015)).

Capitalizing on a different but comprehensive dataset that includes the assignment of refugee immigrants to Denmark from 1986 to 1998, Damm & Dustmann (2014) find that the share of convicted young people in the neighborhood significantly increases both the probability for a male's convictions later in life and the total number of convicted crimes that were executed by men. Their findings suggest that the spillover-effects of neighborhood crime are distinctively linked through the channel of social interaction, which is, however, only true for youth criminal behavior. Because the assignment of the refugee immigrants in their sample was quasi-random, this paper draws on a spatial allocation experiment that does not involve dealing with issues such as endogenous neighborhood selection and thus strengthens the validity of their findings. Bursztyn et al. (2014) used a high-stakes field experiment in Brazil to study different mechanisms of peer effects. In particular, their design allowed them to disentangle two typically confounded channels of social influence in the context of financial decisions, which are social learning (learning from the peer's choice) and social utility (derived utility from the peer possessing the same asset).[18] In terms of economic and statistical significance, their findings indicate that social influence is transmitted through both channels and point in the same direction.

Some studies, however, find little evidence for neighborhood or peer effects. In a highly regarded study, Evans, Oates & Schwab (1992) show that after controlling for selection bias, any measurable peer effect on teenage pregnancy and school dropout rates disappear. While being careful in not claiming that no peer effects exist at all, they rather point critically to methodological issues measuring peer effects in a clean way. Angrist & Lang (2004) use data from the Metropolitan Council for Opportunity (Metco) desegregation program in which mostly black students are sent to more affluent suburbs. Their findings indicate that there are little, if at all, positive spillovers on students. Similarly, Burke & Sass (2013) find little evidence of classroom peer effects on student achievement for Florida public school students. Likewise,

[18] See Bursztyn et al. (2014) for a comprehensive literature review section, in particular on the topic of peer effects in financial settings.

Ludwig & Kling (2007) find little evidence for the contagion of crime hypothesis using MTO data. Instead, their findings indicate that crime rates are merely driven by neighborhood racial segregation.

In conclusion, the existence of peer effects is up for scholarly debate, which is mainly driven by methodological challenges and data problems. However, in following Angrist (2014), the previously discussed MTO program yields the most promising setting to study clean peer effects in the field. Previous research that utilized MTO data yields, among other things, strong gender asymmetries in terms of the evolution of criminal behavior. By common consent, the authors suggest that these findings can be attributed to differences in which males and females respond to their environment and its influences. Ultimately, this leads to differences in magnitude and speed at which (illicit) behavior is picked up.

Evidence from the Lab

In what follows, we shall not attempt to provide an exhausting overview of the comprehensive literature dealing with peer effects in general. Instead, we will focus on a range of influential studies using lab experiments to shed light on mechanisms relating to behavioral spillovers that are more in line with our paper's focus. Later in the paper, we will refer to these studies in more detail where necessary.

Over the last decade, a comprehensive stream of literature studying spillover-effects and behavioral adaptation in the lab has emerged that complement the ongoing important work in the field. As discussed previously, the methodological shift was strongly driven by challenges relating to identifying these effects in a clean way using observational data. Several researchers claim that although the most recent generation of studies measuring such effects with observational data has succeeded to make important steps towards tacking the challenges outlined before, controlled lab experiments are still the gold standard in reducing noise and potential confounds (Angrist, 2014). "However, even if the setting offers an almost perfect opportunity to identify peer effects in many of these studies, the impossibility of controlling for all local or personal confounding factors and for endogenous sorting makes the identification strategy not fully convincing" (Falk & Ichino, 2006, p. 40).

Early laboratory research studying peer effects and social identification jointly has been pioneered by Hoffman, McCabe & Smith (1996) and Bohnet & Frey (1999). These studies made

use of variation in the instruction's wording or enhanced face-to-face communication to study the role of social identification in giving decision, equivocally finding support for its relevance (see also Charness & Gneezy (2008)). We will return to these studies in more detail in chapter 5.

Other studies have looked into peer effects in productivity decisions. In a highly regarded study, Falk & Ichino (2006) found robust evidence for the existence of peer effects in a productivity task. Their results indicate that low-productivity workers are particularly susceptible to peer effects, which results in an over-proportional raise in productivity. Following the work of Mas & Moretti (2009), subsequent studies tried to disentangle the naturally occurring channels of simultaneously observing peers and being observed by peers. For the most part, these studies found the latter channel to be more effective than the former in boosting productivity (cf. Georganas, Tonin & Vlassopolous (2013); for exceptions see Veldhuizen, Oosterbeek & Sonnemans (2014)).

Along the lines of studying behavior in the workplace, Gächter et al. (2012) set up an experiment that investigates reciprocal behavior under observability of other people's actions. They find that the individual's extent to comply with norms of reciprocity is significantly driven by both pay and effort comparison information. Zafar (2011) experimentally examines charitable giving in a social context. He finds that by systematically revealing information, both the learning about descriptive norms (through observing what others do) and the image-related concerns (through revealing own behavior to the reference group) drive individual contribution levels. In a more delinquent context, Falk & Fischbacher (2002) investigate peer effects in the form of conditional stealing behavior. In particular, they investigate whether an individual's inclination to steal is dependent on other peer's stealing behavior. Their main findings suggest that, on the aggregate level, people make stealing decisions conditional on the behavior of their peers.

In economics, a limited number experimental research has also pointed at the contagion of both selfish behavior and dishonesty. Bicchieriy and Xiao (2009) study a dictator game with varying information on other participant's selfish or fair behavior, finding that fairness in actions is contagious. More to the point of our research, Innes & Mitra (2013) use a variant of Gneezy's (2005) deception game to study whether dishonesty breeds dishonesty. Their findings suggest that the beliefs about other's dishonesty is indeed contagious, potentially driven

by the wiggle-room created by such social cues and thus representing a justification device for one's personal dishonest behavior.

4. Drivers of Behavioral Contagion: An Interdisciplinary Perspective

Prior to delving into the subject in more detail, some effort will be made to disentangle many of those existing concepts explaining why and under which circumstances people demonstrate a change in behavior as a function of their peer's behavior. In both economics and psychology, several concepts have been developed over the past few decades referring to similar and often the same reasoning to change one's own behavior.[19] The complexity of the *self* and the dependency of one's own behavior on what one observes of peers has received a lot of scholarly attention (cf. Baumeister (1987), Kahneman & Tversky (2009)). It is thus important to shed light on the underlying concepts driving behavioral adaptation. In this context, we will differentiate between concepts that originated in the (I) economic literature and in the (II) literature of (social) psychology.[20] The economic concepts include (I.1) social decisions and social distance (Akerlof (1997), Glaeser & Scheinkman (2004)), (I.2) image related concerns (Bernheim, 1994), (I.3) taste for conformity (Bernheim, 1994), and (I.4) imitation (Alós-Ferrer & Schlag (2009), Sliwka (2007)). In addition, we also discuss some (social) psychological concepts including (II.1) social learning (Bandura (1971)), (II.2) norms (Cialdini et al. (1990), Bicchieri (2006)), and (II.3) psychological closeness and vicarious dishonesty (Aron & Aron (1986), Goldstein & Cialdini (2007)).

One point is worth clarifying. In this paper, we shall not attempt to settle the argument which scholarly approach explains behavior best. Rather, for our purpose, we will use the previously defined unifying terms *behavioral adaptation* or *behavioral contagion* throughout the paper in order to refer to one person's decision to change initial behavior as a function of observed behavior from at least one other person. Using this umbrella term will allow us to focus on what is relevant without getting lost in conceptual debates. We deem it important to help understand the role social proximity plays in the change of behavior and to what extent proximity mediates the spillover of ethical and unethical behavior.

[19] See discussion in chapter 3.

[20] For a recent discussion of economic and psychological approaches to explaining behavioral adaptation see Bernheim & Exley (2015).

It will be the next sub-chapters' aim to dissect the different approaches in more detail and to explain the reasons why this is the case. At times, the predictions and the empirical outcomes are identical, but the underlying forces causing such outcomes are diverse. In what follows, we will discuss the different theories outlined above and their implications with respect to changes in behavior. At the end of this chapter, we will relate these theories to our experiment and provide a systematic breakdown with respect to the fit of each particular theory to explain behavior in our design.[21]

4.1 Concepts in Economics

4.1.1 Social Decisions and Social Distance

Seminal contributions by Akerlof (1997) and Glaeser & Scheinkman (2004) were among the first to provide an economic framework highlighting the relevance of social distance in affecting social interaction and behavioral adaptation. Beyond rationalizing one's own behavior in an isolated environment, research indicates that behavior is a function of both pure own-maximizing and other-regarding concerns.

For a while, traditional economics has neglected such interdependence and rather put emphasis on individualism. In reality, however, decisions are rarely brought about in total isolation, but are rather the result of an interplay with one's (social) environment. Arguably, individuals care about both status concerns (in absolute and relative comparisons) and other's well-being (Fehr & Schmidt (1999), Bolton & Ockenfels (2000), Charness & Rabin (2002)). Because social interactions typically render externalities with the potential of slowing down conversion towards socially (un-)desirable equilibria, it is important to understand the underlying mechanism.

In his attempt to explain the connection between social interaction and behavioral adaptation (conformity, in particular), Akerlof (1997) made use of the Newtonian theory of gravity. Akerlof's approach centers on explaining conformist behavior as a function of distance in the social space. Such difference in one's social space is characterized by the difference in behavior between oneself and a person or group of relevance. For his purpose, Akerlof applies the concepts of a gravity model allowing him to argue that conformity leads to benefits (e.g. higher

[21] It is worth noting, however, that our discussion here will be purely descriptively. We will contrast the explanatory power of these concepts within the frame of our experiment in chapter 6.

individual utility or gains from trade) that are negatively correlated with distance in social space.[22] Akerlof puts a reduced form concept forward in which one's utility declines as the individual's behavior deviates from the behavior of others. Assuming the existence of representative agents, his model predictions yield a set of equilibria in which the ultimate behavior of all individuals is the same, thus clearly characterizing the behavior of every party.[23]

Ultimately, in order to maximize one's intrinsic utility, the individual will converge towards the observed behavior and thus reduces the existing social distance.[24] With reduced distance, social interactions become more favorable through facilitation of e.g. mutually beneficial trade.

4.1.2 Image Related Concerns

In understanding individual behavior, one fundamental question arises: "How much of what we do is the result of our genetic code (nature); and how much of it is a function of our environment, including the actions of those around us and our own past actions (nurture)?" (Cabral, 2005, p. 15). The threat of reputation loss and being punished in the case of wrongdoing is an integral component of individual decision making in general. By weighing costs against benefits, the impact of possible reputation loss might deter individuals to engage in illicit behavior of any kind.

Individuals are striving for social inclusion and recognition. They want to belong to one or more social groups and engage in social interaction. Striving for social identity is an inherent characteristic, thus decisively driving the individual's yearning for maintaining an adequate

[22] Because geographic location is one determinant of social interaction, this approach is a generalization of what sociologists have coined "social geography", which has been inspired by Krugman's (1991) work on economic geography.

[23] Although *ex-post* behavior is uniform across individuals in equilibrium, such an approach still models conformism because of the *ex-ante* desire to be and behave like others. Akerlof also proposes an extension with individual heterogeneity that allows for the formation of social sub-groups with own norms and values. The extension models a mechanism in which (randomly distributed) past social location for each individual is inherited. In combination with static expectations about "the positions to be occupied by others in social space […] such a model can portray stable groups in low level equilibrium traps because individual's incentives to choose x to conform with those whose inherited social locations are close may overwhelm their incentives to choose x for intrinsic reasons" (Akerlof, 1997, p. 1010). In our experiment, we will study this mechanism by holding the other's behavior fixed and thus allowing for deliberate behavioral adaptation free of unstable, higher-order, or even nonexistent beliefs about the other person's next move. We will return to this argument in our design section.

[24] A point not made clear in this approach is whether such desire to conform is also dependent on the signaling value of one's behavioral adaptation. The model argues that the desire to conform and reduce social distance is entirely driven by intrinsic concerns. However, the inability to signal conformity to one's peers runs counter to the initial concept of conformism (introduced by, for example, Ash (1958) and Bernheim (1994). Thus, this concept rather resembles the ideas of imitation of preferences (Sliwka, 2007) or self-expansion (Aron & Aron, 1986) that will be addressed in the next sub-chapters.

reputation. In this sense, losing face as the result of deviant behavior may serve as an avoidance, as this might result in exclusion from the group (cf. Bernheim (1994)). Along these lines, Akerlof (1980) argues that the deviation from social standards might lead to loss of social reputation and consequently to ostracism. However, it is worth noting that *deviant behavior* not necessarily implies bad behavior per se, but has to be understood as a deviation from the underlying social norm in either direction. Reputation can also work as a means to preserve a coherent self-image that allows keeping an internal consistency (cf. Baumeister (1998)). The concept of the self has multiple facets and impact behavior in different ways, as level of self-regard and self-esteem, the extent and content of self-identity as well as the structure of the self-concept exhibit motivational implications (cf. Wells (1978)). By that, people try to avoid a negative update of their self-image through their actions, which otherwise would result in an internal conflict. Such a mechanism possesses the power to prevent individuals from engaging in delinquent behavior in the first place. This mechanism is more distinct when the group is salient, consequently giving rise to more extensive alignment with the group's behavior. What is more, own actions could also be subject to social signaling, which highlights the perception of oneself by others and thus potentially triggering conform behavior (cf. Grossman (2010)). It is important to clearly distinguish between these two motives and mechanisms when analyzing one's individual drivers for conform behavior.[25]

However, for image related concerns to be effective, a setting of repeated interactions has to be in place. In a one-shot setting, the threat of a reputation loss is negligible as the players won't face each other for a second time. If recurring interaction is not taking place, any deviant behavior can neither be traced back to the individual nor will be colored negatively with regard to future collaboration. Existing research supports this perspective. For example, there is comprehensive experimental literature pointing to the fact that contributions and compliance with underlying norms rise in multiple-shot interactions. In particular, this is true when transparency (e.g. ability to observe other participants' contribution decisions) or punishment (e.g. for non-compliance with social norms) is possible (cf. Andreoni (1995), Fehr & Gächter (2000), Cameron et al. (2009), Chaudhuri (2011)). Consequently, in order for reputation to be effective (even in absence of a punishment-mechanism), a setting of repeated interaction has to be in place and actions have to be observable to other people.

[25] These aspects will be taken into consideration in the experimental design. We will return to this point later.

4.1.3 Taste for Conformity

Research and observations of real life situations indicate that underlying social factors decisively influence individual behavior. Motivation for a particular behavior is driven by the inherent desire to be valued and to win prestige, esteem, popularity, and acceptance (Ellingsen & Johannesson, 2008). Arguably, both the introversive coherence and outward appreciation matters and can be achieved through conformity in behavior. The basic idea here is that any infinitesimal deviation from effective norms might be punished by the social group (Bernheim (1994); see also Akerlof (1980) and Tirole (1996) for related work on the role of reputation in individual decision making and behavioral alignment).

In his seminal contribution, Bernheim (1994) formalized the concept of taste conformity. He argues that the individual's preference for status is in line with psychological, evolutionary, and behavioral considerations. In particular, natural selection favors concerns for status as this goes hand in hand with greater opportunities for reproduction. From a behavioral point of view, such concerns act as a reinforcing device to form preferences for higher esteem because esteemed individuals are more likely to receive better treatment.

Although partly representing a departure from the traditional formulation of preferences, this approach does not necessarily require the abandonment of consistent, self-interested optimization processes. By assuming that status depends on "public perceptions about an individual's preferences over actions […] esteem is determined by expectations about future actions and that tastes and proclivities are the best predictors of future actions" (Bernheim, 1994, p. 843).

Unsurprisingly, research indicates that people share similar preferences and opinions within the same group, which can arguably be both the antecedent and the effect of social interaction and norm alignment.[26] Aligning with existing norms is a way to signal one's (wishful) belonging to a certain social group. As argued by Hogg & Tindale (2002), enacting in-group-prototypical behavior is a way to validate one's own group membership, not only to the social group but also to themselves. By that, aligned behavior might function as a signal to both the external (e.g. the peers) and the internal world (e.g. themselves) in order to solidify the sense of belonging. If the individual cares about status, one has to set the right signals to the peers

[26] Manski (1993) refers to this as the previously mentioned reflection problem, rendering it extremely difficult to study clean peer-effects in the field. We will return to this point shortly in the experiment's design section of chapter 6.

in order to create esteem. This esteem will be provided in case the individual behaves in a way that is expected from him.

In this vein, Bernheim (1994, p. 844) derives the following proposition, highlighting the decision process leading to behavioral adaptation:

> *"When popularity is sufficiently important relative to intrinsic utility (defined as the utility derived directly from consumption), many individuals conform to a single, homogeneous standard of behavior, despite heterogeneous underlying preferences. They are willing to suppress their individuality and conform to the social norm because they recognize that even small departures from the norm will seriously impair their popularity."*

Arguably, the extent at which people are willing to align with their peer's behavior is driven by the degree of social identification to the social group. Social identity theory provides helpful guidance to distinguish between different magnitudes of personal identification with the peer's behavior (cf. Tajfel & Turner (1979), Tajfel (1982)). According to this approach, people categorize themselves and other persons into different social groups. Membership in a group is defined as the social identity and individuals strive to enhance their position and self-esteem through actions. These actions encompass the alignment with existing group norms. "[…] the social identity analysis of categorization processes suggests that group cohesion or solidarity is not only attraction among group members, but also attitudinal and behavioral consensus, ethnocentrism, in-group favoritism and intergroup differentiation, and so forth – the entire range of effects of categorization-based depersonalization" (Hogg & Tindale, 2002, p. 65). A comprehensive stream of literature suggests that such identification is driven by in-group-out-group concerns, finding that in-group favoritism is found across different contexts and social settings (cf. Tajfel (1982), Hoffman, McCabe & Smith (1996), Bohnet & Frey (1999), Eckel & Petrie (2011), Charness & Gneezy (2008)). From this one can derive the assumption that the stronger the desire to be part of the in-group, the stronger one's intrinsic willingness to comply with prevalent rules and engage in conformity by adapting social norms. Such a desire is driven by, for example, the inherent relevance of the social image component (cf. Andreoni & Bernheim (2009)). The stronger the social identity with a reference group, the stronger the effect of social comparison on individual conformity. Overall, taste for conformity is a function

of existing social identity and the extent of social comparison to one's reference group. Deviating from the group's behavior (even in extreme cases where the group behavior is clearly wrong, cf. Asch (1958)) and thus being perceived as different might create discomfort in form of substantial disutility.

4.1.4 Imitation

In this subchapter, we will discuss two streams of literature suggesting that behavioral adaptation is the result of either pure behavioral imitation (e.g. due to observing behavior that leads to a superior outcome) or the adjustment of preferences (e.g. contingent on the behavior of the peers). With no claim to completeness, this chapter's focus will lie on the selected sample of contributions discussing the topics of behavioral imitation and adjustment of preference as a mechanism for behavioral adaptation.

Pure Behavioral Imitation

Imitation is a common behavioral trait of both animals and humans (Laland, 2002). While it is likely that imitation is driven by evolutionary and cultural facets, its pervasiveness can only be explained if it sufficiently often leads to a desirable outcome. Imitation has a number of desirable features because it allows one, among other things, to *free ride on superior information* (in the spirit of social learning, but see also Sinclair (1990)), to *save mental resources* (in the spirit of bounded rationality, but see also Conlisk (1980)), or as a means *to be regarded similar to others* (in the spirit of the social esteem argument, but see also Cho & Kreps (1987)).

Along these lines, Alós-Ferrer & Schlag (2009) abstract from the standard Bayesian belief-based approach and introduce the concept of imitation as a belief-free behavioral rule. The authors make the case that imitation is triggered by the observation of a superior outcome. In face of different imitation concepts, the degree of sophistication involved in assessing the extent of the observed agent's better performance clearly dominates a simple 'follow whoever performs better than me' rule. The reason behind this argumentation is that "the reluctance to switch when the observed choices are only slightly better than the own might not be due to switching costs, but rather to the fact that the payoff-sensitivity of the imitation rule allows the population to learn the best option" (Alos-Ferrer & Schlag, 2009, p. 273). In essence, this approach highlights the idea that behavioral adaptation can be triggered by behavioral imitation that results from observing (sufficiently) superior outcomes.

Adjustment of Preferences

Another strain of literature models behavioral adaptation in a way that is hard to swallow by traditional economists: preferences are not stable per se but are subject to interrelations with the encompassing ecological system (cf. Sliwka (2007)). In his work, Sliwka (2007) introduced a third group of agents (in addition to agents of the homo oeconomicus type and agents with other-regarding preferences, cf. Fehr & Schmidt (1999), Bolton & Ockenfels (2000)), namely the 'conformists'. Here, such conformists are stochastic learners and exert behavioral adaptation through imitation of preferences.

Such conformists find themselves in a situation of uncertainty concerning the appropriate and acknowledged behavior by the social group. By such definition, a conformist will then be inclined to adapt his own behavior if he believes that this particular behavior is represented, meaning sufficiently supported and performed, by other agents. Observing certain behavior induces learning about other peoples' preferences, which in turn potentially triggers behavioral adaptation on the side of the agent.

Arguably, a conformist of the Sliwka-type is an agent who is driven by moral convictions and, although being mindful of social norms, will only be guided by them if he believes that a sufficiently large subset of the relevant social group will behave in the same way. He will also suffer remorse only if such feelings of guilt are sufficiently represented within the social group (also see Kandel & Lazear (1992)).

While classical economists might find the idea of contingent preferences hard to swallow, the recent years' literature has steadily evolved modeling social preferences beyond what has typically been assumed to be the case for fully rational agents. This approach is backed up by a variety of findings indicating that individual taste is not invariant but merely subject to social influence (Salganik, et al., 2006), which is in line with existing psychological literature on the pertinence of social influence (cf. Sherif (1937), Ash (1952), and Cialdini & Goldstein (2004)).

4.2 Concepts in (Social) Psychology

4.2.1 Social Learning Theory

Individuals constantly send signals through their behavior. These signals soak through different channels and as such not only reveal personal preferences but also transmit information

about the current state-of-the-world (Anderson & Holt, 1997). Observed behavior is constantly evaluated by the environment (i.e. the peers), triggering an action that is frequently referred to as *social learning*. Potentially, this allows other individuals to infer relevant information with respect to *accepted* or *correct* behavior (cf. Banerjee (1992), Ellison & Fudenberg (1993), Bikhchandani et al. (1998), Bikchandani & Sharma (2000), Chari & Kehoe (2004)). Broadly speaking, this particular strain of literature suggests that behavioral adaptation can be subdivided into two fundamental mechanisms, one of which is *social learning* and the other, *social utility* (Bursztyn, et al., 2014)).

Along these lines, Bandura's (1986) Theory of Social Learning emphasizes that behavior is the result of what has been learned through the observation of others rather than through trial-and-error. In particular, social learning is referred to in situations where individuals infer from a third party's behavior, such as buying decision, that the product must be of high quality. This in turn initiates own buying decisions. With respect to financial markets, herding behavior based on social learning assumes that conscious actions, such as buying decisions of an asset, transmit a positive signal about the product's value to the observer, thus increasing the probability that this action will be imitated. Under social learning, individuals constantly engage in a form of Bayesian-updating and adjust their understanding of the world.[27] The current trend of *social trading* notices this development, which utilizes the power of swarm intelligence to carry out investment decisions on the stock market (cf. Neumann et al. (2013)). By following the crowd, these people anticipate some existing information asymmetry and assume that the ones who act are the ones who are better-informed (cf. Conlisk (1980), Bernheim (1994)).

In sum, Bandura's approach stresses the idea that individual behavior is driven by informational motives and that we learn from direct experience, interaction, and observation. Consequently, social interaction allows individuals to learn and adapt established behavior.

4.2.2 Norms

By nature, humans are social creatures who use norms as guidance to learn what should and should not be done. The social context shapes the understanding and perception of existing

[27] Conversely, *social utility* implies that "one's utility from possessing an asset (or product) depends directly on the possession of that asset (or product) by another individual." (Bursztyn, et al., 2014, p. 2). Although of particular relevance for explaining behavioral traits on, among others, financial markets, this approach will not be investigated further.

norms, which in turn substantially affect individual behavior. Along these lines, the contributions by Cialdini et al. (1990) and Bicchieri (2006) have fueled our understanding of norms in general and their role in shaping social interactions.

We start by noting that research is disunited on what exactly is meant by a social norm. In addition, Postlewaite (2011) argues that economists have often been reluctant to incorporate social aspects such as norms into their analysis, especially if the results lead to a deviation from the standard neoclassical predictions. This, in turn, triggers even more fragmentation within and among the fields of research. For our purpose, we define norms as "the result of shared notions of appropriate behavior and the willingness of individuals to reward appropriate behavior and punish inappropriate behavior. […] As long as the rewards and punishments are sufficiently large, norms can stabilize a vast range of different behaviors" (Boyd & Richerson, 2001, p. 283). *Appropriate behavior* is believed to be represented by a social norm if "the behavior differs from that of other groups in similar environments" (Postlewaite, 2011, p. 32). From the perspective of game theory, Binmore (1998) adds that social life is an infinite game with norms representing equilibria that are in line with the folk theorem where both types of individuals are punished, norm violators and those who fail to punish them.[28] From this, we can conclude that social norms sustain because individuals have, among others, "a desire to coordinate, [a] fear of being sanctioned, [and want to] signal membership in the group, or simply follow the lead of others" (Young, 2015, p. 359).

In their seminal work, Cialdini et al. (1990) separate two types of norms that drive individual behavior. On the one hand, the *descriptive norms* induce an understanding of what is (supposedly) *normal* and thus mirrors the status-quo. Observing other people's behavior triggers a decisional shortcut as one relies on what seems to be sensible based on observed behavior. Conversely, the *injunctive norm* mirrors an individual's beliefs about how someone (supposedly) *ought* to behave. This norm is socially loaded and highlights what is morally approved or disapproved. This said, "rather than simply informing one's actions, these norms enjoin it

[28] Boyd & Richerson (2001) provide a theoretical model explaining the mechanism of norm adaptation under the existence of occasional learning and inherent conformism. For an extended evolutionary perspective see McAdams (1997), Bowles & Gintis (2004), and Richerson & Boyd (1998). However, the question of why individuals choose to enforce norms remains open. We follow Boyd & Richerson (2001, p. 283) and assume "that the problem of why people choose to enforce norms has somehow been resolved […] over the long history of human evolution."

through the promise of social sanctions. Because what is approved is often what is typically done, it is easy to confuse these two meanings of norms." (Cialdini, et al., 1990, p. 1015).

In addition, Bicchieri (2006) introduces the notion of conditional preferences (preferences that are conditional on social expectations) and unconditional preferences (preferences that are purely inherent to the individual, like self- and other-regarding preferences). In terms of individual norm compliance, Bicchieri argues that expectations are the main drivers. Here, the individual's inclination to obey norms is driven by both social beliefs, which are represented by normative and empirical expectations, as well as by non-social beliefs, such as factual and personal normative beliefs. Social beliefs are represented by two sets of expectations. For one, *normative expectations* describe second order beliefs about what others think should be done. For another, *empirical expectations* represent what one expects others will do. As elaborated in Bicchieri & Xiao (2009), such expectation can be grounded on past observations of conformity or its consequences. Individual conformity is thus conditional on both belief sets. In turn, non-social beliefs are represented by *factual beliefs* (that is, beliefs about the actual consequences of behavior) and *personal normative beliefs* (that is, what one personally thinks should be done). "A social norm then is a collective practice sustained by empirical and normative expectations and by preferences conditional on both these expectations" (Bicchieri, et al., 2014, p. 4).

Norms play an important role in everyday life and are decisive in establishing and sustaining social interactions. They determine the way we perceive the world and make sense out of it. The existence of norms provides an explanation to why and when behavior is contagious since individuals often follow and subordinate to the prohibition, indulgence, or encouragement of (un)ethical norms without much deliberation or putting it into question (Boyd & Richerson, 2001). In addition, Boyd & Richerson (2001) suggest that the saliency of norms depends on its source (e.g. on whom one interacts with), thus suggesting that norm obedience is likely to be a function of culture in general and social proximity in particular. This sometimes leads to the perverse outcome of antisocial punishment, in which social behavior is punished because it does not correspond to the society's norms (e.g. Herrmann, Thöni & Gächter (2008)).

4.2.3 Psychological Closeness and Vicarious Dishonesty

In social psychology, there is a comprehensive stream of literature introducing concepts of psychological closeness and vicarious dishonesty. These concepts serve the purpose of explaining the mechanism of behavioral adaptation in various contexts. A brief discussion of the more prominent approaches follows.

The self-expansion theory as proposed by Aron & Aron (1986) and extended by Aron et al. (1992) argues that the underlying mechanism of close relationships is defined by the inclusion of others in the self. In particular, the broadening of one's self is likely to occur if we are (or at least feel) close to the other person. Existing research highlights a number of factors that influence one's feelings of attachment towards other persons. Among these, sharing a common group membership (Tajfel, 1982), a similar name (Pelham, et al., 2005), the same birthday (Cialdini & DeNichols, 1989), or even imagining closeness (Goldstein & Cialdini, 2007) are found to trigger self-expansion and thus matter for conformity to selfish or unfaithful behavior (Gino & Galinsky, 2012). Hence, individuals take the perspective of others, afford them "self" status, and are likely to adopt other people's behavior when social closeness is imagined (Galinsky, et al., 2005). Along these lines, Goldstein and Cialdini (2007) argue that when closeness is imagined, individuals make inferences about the other person's attributes based on observed behavior and reflect upon them just as if they would have engaged in the very same behavior themselves.

As psychological closeness blurs the gap between self and others, such vicarious self-perception processes facilitate mimicry and render behavioral adaptation more likely the stronger the feelings of attachment are. In this vein, Gino and Galinsky (2012) argue and find experimental support that psychological closeness to an unethically acting person creates distance from one's own moral compass, thus increasing the likelihood of conformity towards unethical behavior.[29]

[29] More concepts emphasizing the degree of social distance between individuals as a driving factor of behavioral adaptation are discussed in more detail in chapter 5.1.

5. The Role of Social Identity and Observed (Un)Ethicality in Behavioral Contagion

5.1 Conceptual Introduction

Beyond Becker's (1968) seminal work on the rational criminal, recent work within the field of economics started implementing factors such as identity and culture into an individual's decision framework (Akerlof & Kranton, 2000). While this stream of literature does not necessarily yield hypotheses contradicting those generated by the rational crime economics literature, the inclusion of identity, morals, and reputation allows one to draw a more comprehensive picture of individual decision-making (Bénabou & Tirole, 2011).

In fact, existing research points to the relevance of social identity in determining social interaction such as the degree of trust and reciprocity (Leider, et al., 2009). Psychologists have devoted decades of research efforts to the study of group identity, which is closely related to the concept of social identity. Group identity is a function of social identity and is driven by one's identification with a social group. Sufficient identification triggers an in-group bias, indicating a preference for one's identified in-group along different dimensions (see also Gino, Ayal & Ariely (2009)).[30] Exemplarily, Chen & Li (2009) use painting preferences as a social identity and a matching device to study charitable giving, envy, punishment, and social concerns. Some research has also investigated the role of social identity for cooperation (Eckel & Grossman, 2005) and equilibrium selection (Chen & Chen, 2011).

Beyond the concepts explaining general behavioral adaptation that we have already discussed in chapter 3, some theories predict that social identity affects the magnitude of behavioral adaptation. Exemplarily, McLeod et al. (1966) introduced the concept of the *attraction effect*, arguing that conformity is mediated by the level of social identification between two people or a group. The *self-expansion theory* highlights that the people's sense of the *self* can be broadened in order to include others and even more so if we feel close to them. In consequence, the person's self-bias (e.g. egoism) includes not only himself but also extends to others (Aron & Aron, 1986). Within the framework of the *vicarious self-perception theory*, observing the behavior of people that *are closer* from a social identity perspective may induce a carry-over effect of behavior (Goldstein & Cialdini, 2007), whereas the *perspective-taking theory*

[30] See Chen & Li (2009) for a comprehensive review of related literature.

assumes that the *imagined* social closeness is already sufficient to facilitate behavioral adaptation (Galinsky, et al., 2005).

As has been discussed previously, psychological research has introduced a number of ways to proxy social identity in the lab. Due to methodological differences, existing economic research has resorted primarily to using rough proxies to introduce social identification in one form or another. Exemplarily, this has been achieved by varying the wording of the experimental instructions (Hoffman, et al., 1996), using face-to-face interaction (Bohnet & Frey, 1999) or showing pictures (Eckel & Petrie, 2011), revealing names (Charness & Gneezy, 2008), revealing preferences such as those for paintings (Chen & Li, 2009), and recruiting friends and family members (Brandts & Solà, 2010). In line with most psychological research, these studies find that social identity triggers in-group favoritism.

However, two natural problems arise with these approaches to introduce social identification in the lab: firstly, the introduction of potential biases caused by face-to-face communication within the lab or letting participants interact with their kin (Roth, 1995). Exemplarily, it has been shown that individuals discriminate using social cues such as looks or skin color (Eckel & Petrie, 2011). Additional research has investigated other-regarding behavior as a function of social distance. Using the dictator game setting, Bohnet & Frey (1999) find that other-regarding behavior is more pronounced under stronger identification. They incepted stronger identification by the use of identification prior to the dictator game. Again, their approach is only a rough proxy for social proximity and does not control for potential biases stemming from, among others, unobserved attraction effects. Secondly, and more along the lines of our experiment, applying these methods to introduce social identification does not usually allow for a variation of social proximity beyond a simple binary outcome of no social identification versus some social identification. In our experiment, we propose a method that allows us to naturally vary the degree of social identification in the lab without having to deal with the issues raised above. We will return to this in our design section.

Overall, it may be stated that different streams of existing research point at the significance of social identity in explaining the magnitude of behavioral adaptation thus making it worthwhile to examine this topic in more detail.

The ethicality of observed behavior might also be decisive to the magnitude of behavioral adaptation. As we will argue throughout this paper and will be explained more formally in chapter 5.2.2, we expect that observations of unethical behavior render behavioral adaptation more likely and more extensive as compared to observations of ethical behavior. Existing empirical literature on the slippery-slope effect (cf. Gino & Bazerman (2009), Welsh et al. (2015)) as well as on the "broken windows effect" (cf. Beckenkamp et al. (2014), Lefebvre, Pestieau, Riedl & Villeval (2015)) support this point of view.[31] Arguably, the magnitude of behavioral adaptation is asymmetric in one direction or another, indicating that the likelihoods for adapting behavior are different when observing ethical or unethical behavior. Below, we will provide some theoretical arguments for this claim before turning to a more formal approach.

Individuals engage in ethical and unethical behavior for all sorts of reasons and sometimes turn evil extremely quickly under particular circumstances (see the infamous Milgram experiment (Milgram, 1963) and Zimbardo's Stanford prison experiment (Zimbardo, 1971)). Different perspectives are helpful in understanding one's rationalization to engage in (un)ethical behavior, such as self-rationalization, obedience to norms, or active signaling to one's peers.

Individuals tend to rationalize their behavior with themselves while trying to avoid cognitive strain stemming from, among others, negative self-image updating (cf. Wells (1978), Baumeister (1998)). Here, individuals do not want to think of themselves as a bad person and thus tend to engage in self-deception and moral rationalization and create positive disillusions that sustain a positive self-image (cf. Taylor (1989), Mazar, Amir & Ariely (2008)). "Self-deception allows one to behave self-interestedly while, at the same time, falsely believing that one's moral principles were upheld. The end result of this internal con game is that the ethical aspects of the decision "fade" into the background, the moral implications obscured" (Tenbrunsel & Messick, 2004, p. 223). Arguably, by avoiding thinking of immoral acts as such, mental processes leading to a denial of facts facilitate a person's inclination towards unethical behavior.

From the perspective of norm theory, the exposition of unethical behavior comes with a relatively higher risk of being condemned by and excluded from the social group (Bicchieri, 2006). The degree to which one can behave unethically (if at all) among one's social peers is typically

[31] Although not explicitly along the lines of (un)ethical behavior, a similar argument can be made based on the findings within the extensive literature on conditional cooperation and punishment (see Chaudhuri (2010) for a selective review).

grounded in uncertainty. Consequently, a person's decision to engage in either ethical or unethical behavior is not only simple cost-gain maximization under risk but also involves accounting for the inherent uncertainty. The uncertainty involves, among others, the deviation from what is deemed appropriate within the social group, reputation loss, and the potential exclusion from the group. From a psychological perspective, engaging in unethical behavior involves higher costs, as one has to forcefully convince oneself that gaining personally at the expense of harming others is acceptable. Processes leading up to this reasoning involve, among others, self-deception, justification, and hypocrisy. Typically, only the acting individual profits from unethical behavior in (non-) monetary terms, such as stealing money, while another person is losing out. In turn, however, normally both individuals gain from ethical behavior such as donating money: the receiving individual gains in monetary terms and the acting individual, while losing out in monetary terms, is (over-)compensated from a behavioral perspective by either the positive feelings of warm glow of giving and altruism (Becker (1974), Andreoni (1989) & (1990)) or social pressure (Akerlof & Kranton, 2000). Plausibly, to convince oneself to behave unethically involves more (psychological) effort than to behave ethically.

From a signaling perspective, observing someone who is socially closer to oneself as compared to observing a stranger sends a more salient form of social learning about appropriate (or at least tolerated) behavior within the social group. This assumption is also along the lines of Schelling's (1968) prominently stated "the more we know, the more we care" and a number of experimental results in various settings (see Eckel & Grossman (1996), Bohnet & Frey (1999), Charness & Gneezy (2008), and Gino & Gallinsky (2012)). The resulting resolution of uncertainty allows the individual to overrule their own concerns about the inappropriateness of unethical behavior. This stream of literature suggests that observing unethical behavior might trigger stronger contagion as compared to observing ethical behavior. For these and more reasons, we expect behavioral adaptation to be asymmetrically biased towards unethical behavior.

5.2 Theoretical Model

We extend the model introduced by Akerlof (1997) and expand the underlying concepts in a way that is conducive to understanding the role of social identity and the (un)ethicality of

observed behavior in behavioral contagion.[32] Previous work put emphasis on modeling social interactions and behavioral adaptation as a function of distance in the social space, e.g. resulting from distance in behavior and geographic location to one's peers. We, however, shed light on the relevance of social distance and proximity resulting from *overlapping and common interests and preferences* in driving behavioral adaptation. Intuitively, one could reasonably assume that the behavior of one's family and friends exhibits a more salient signal and thus is taken more strongly into consideration than observing random strangers. What is more, we follow the previous discussion and assume additionally an asymmetry in behavioral contagion depending on the ethicality of the observed behavior. This is a substantial extension to Akerlof (1997), where an asymmetric adaptation in behavior cannot be illustrated.

The purpose of this section is to outline a simple reduced-form model that allows one to draw predictions about individual contribution decisions in the various contexts that are reflected by the experimental design. These predictions can then be tested empirically. The principal goal of this exercise is to show how behavioral contagion is mediated by social identity and the (un)ethicality of observed behavior and how these factors affect the *adaptation gap*, that is the difference between one's own and observed behavior, after contagion has taken place. We will introduce the symmetric model first before turning to the extension.

5.2.1 Symmetric Contagion

Prior to introducing the formal model, we begin with the intuition of what the model should capture. The underlying idea is that individual behavior is a function of the peer's actions and thus encompasses more than one's self-referentiality. In a situation of social interaction, an individual is expected to face a trade-off, weighing one's own preferences against the peer's revealed preferences. To stay in Manski's (1993) terms, social interaction and the understanding that one's own actions are reflected by the peers and thus have an impact on other (closely related) people shapes the individual's willingness to engage in behavioral adaptation. The

[32] In particular, we focus on Akerlof's (1997) quadratic utility version of the conformity model as this generates unique equilibria from which we can derive testable hypotheses for our experiment, which we also deem to be more in line with the general story of this paper. While Akerlof's general model puts more emphasis on one's own decisions, the quadratic utility model has an a priori assumption that puts one's own and peer behavior on an equal footing with respect to how behavior affects individual utility. Assuming that individuals are on a continuum between extremely selfish and extremely altruistic, the quadratic utility approach is conclusive. It should be noted that this assumption is not crucial to our model's predictions as we mainly focus on the relevance of two factors, i.e. social identity and the (un)ethicality of observed behavior, in affecting one's own decisions. Thus, the equal-weight assumption is not decisive in predicting the direction in which individual behavior changes as a function of those two factors.

magnitude to which one is willing to revise one's own initial behavior and adapt is first and foremost a function of the social proximity to the peer. "As a consequence, the impact of my choices on my interactions with other members of my social network may be the primary determinant of my decision, with the ordinary determinants of choice the direct additions and subtractions from utility due to the choice) of only secondary importance" (Akerlof, 1997, pp. 1006-1007). It is thus reasonable to assume that the individual's utility is subject to a relative evaluation of one's own behavior and the behavior of the peers. Similar to Akerlof (1997), the underlying characteristic of our model is the feature that individual utility is declining with increasing distance between one's own behavior and the peer's behavior. While the aim of this section is to outline a model that explains changes in general behavior, we will use the resulting predictions to generate hypotheses in the context of (un)ethical behavior.

In our design, we resort to a two-stage dictator game in which each individual is paired with a charity of his choice. Both, the individual and the charity start with an initial endowment $\boldsymbol{I_c} = \boldsymbol{I_i} = \boldsymbol{I} \in \mathbb{R}^+$ of equal size. At each stage $\boldsymbol{t} \in [1, 2]$, each individual $\boldsymbol{i}$ faces the choice, x_t^i, of either (a) donate (part or all of) one's own money to the charity, (b) retain the equal split, or (c) take away (part or all of) the charity's money and add it to one's own income. We will refer to (a) and (c) as ethical or unethical behavior, respectively. Naturally, the individual's decision is of the form $\boldsymbol{x_t^i} \in [-\mathrm{I}, +\mathrm{I}\,]$. The only difference between both stages is the information set that the individual possesses about his peer's behavior. That is, after completion of stage 1, the individual observes a random individual's behavior from stage 1. At stage 2 (that is, after the observation), the individual is given the opportunity to revise his initial decision, if desired.

Let $\boldsymbol{\alpha_t} \in (0,1)$ depict the social proximity of an individual at time $\boldsymbol{t}$. Importantly, let an individual's inherent attitude towards (un)ethical behavior be described by θ^i. That is, θ^i represents the individual $\boldsymbol{i}$'s preference to give or take a particular monetary amount within the boundaries of one's income in a given situation, thus being defined as $\boldsymbol{\theta^i} \in [-\mathrm{I}, +\mathrm{I}\,]$. What is more, let ψ_t^{ij} represent individual $\boldsymbol{i}'s$ prior (stage 1) or actual observation (stage 2) of individual $\boldsymbol{j}'s$ (un)ethical behavior.

$$\boldsymbol{\psi_t^{ij}} = \begin{cases} \widetilde{x_1^j} & if\ t = 1 \quad (truncated\ normally\ distributed\ prior) \\ x_1^j & else \qquad\qquad\qquad\quad (updated\ beliefs) \end{cases} \tag{1}$$

From this it follows that each player maximizes own utility at Stage 1 of the following form:

(2) $$\max_{x_t^i} U_t^i = I - \left(1 - \alpha_t\ \right) * \left(x_t^i - \theta^i\right)^2 - \alpha_t\ * \left(x_t^i - \psi_t^{ij}\right)^2$$

At stage 1, individual $\boldsymbol{i}$ has no information about individual $\boldsymbol{j}$'s decision, as all participants carry out their decisions simultaneously. Consequently, $\boldsymbol{i}$ resorts to forming beliefs about the behavior of $\boldsymbol{j}$. As depicted above, individuals face a trade-off decision at stage 1, in which deviation from the individual inherent characteristic θ^i has to be weighed against deviating from one's own beliefs about the peer's behavior ψ_t^{ij}. Because no information about the peers was given at stage 1, the resulting decision is a simple maximization problem of the form:

(3) $$\frac{\partial U_t^i}{\partial x_t^i} = 2\alpha_t\ * \left(\psi_t^i - x_t^i\right) - 2 * \left(\alpha_t\ - 1\right) * \left(\theta^i - x_t^i\right)$$

$$\Leftrightarrow x^* = x_t^i = \theta^i - \alpha_t\ \theta^i + \alpha_t\ \psi_t^{ij}$$

yielding the comparative static:

(4) $$\frac{\partial x_t^i}{\partial \alpha_t} = \psi_t^{ij} - \theta^i$$

We can infer that a change of $\boldsymbol{x_t^i}$ in $\boldsymbol{t}$ depends on the social proximity (that is either a prior in $\boldsymbol{t = 1}$ or an updated belief in $\boldsymbol{t = 2}$**)** in the following way:

(5) $$\frac{\partial x_t^i}{\partial \alpha_t} = \begin{cases} > 0 & if\ \psi_t^{ij} > \theta^i \qquad (x_t^i\ increases\ in\ \alpha_t\ ) \\ = 0 & if\ \psi_t^{ij} = \theta^i \qquad (x_t^i\ unaffected\ by\ \alpha_t\ ) \\ < 0 & if\ \psi_t^{ij} < \theta^i \qquad (x_t^i\ decreases\ in\ \alpha_t\ ) \end{cases}$$

In order to study the adaptation gap at stage 2, that is the difference in behavior between individual $\boldsymbol{i}$ and $\boldsymbol{j}$ after the second stage, one has to hold $\boldsymbol{j}$'s behavior from stage 1 constant while giving individual $\boldsymbol{i}$ the ability to revise his initial decision.[33]

Proposition: Equation (2) provides a solution to the maximization problem and reduces the adaptation gap to:

(6) $$\left|x_2^i - \psi_2^{ij}\right| = \left|\theta^i - \alpha_2\ \theta^i + \alpha_2\ \psi_2^i - \psi_2^i\right| = \left|\theta^i * \left(1 - \alpha_2\ \right) + \psi_2^i(\alpha_2 - 1)\right|$$

[33] Which is exactly what we will do in our experiment. See chapter 6.

Observe that unlike in Akerlof's (1997) general conformity model, this approach generates a unique equilibrium prediction due to restrictions put on the social proximity parameter, that is $0 < \boldsymbol{\alpha_t} < \mathbf{1}$ and the linear reaction function.

Essentially, by assuming that behavioral adaptation is symmetric in either direction this indicates that the gap is driven by the social proximity to the peer that one is observing: the closer individuals are in terms of proximity the smaller is the expected gap and the more similar is their behavior. An alternative interpretation is that the peer effect is stronger and leads to a more extensive behavioral adaptation the higher their social proximity is.

5.2.2 Asymmetric Contagion

The main purpose of the extended model is to allow for an asymmetric adaptation to unethical and ethical behavior. It is plausible to assume that not only observing but also starting to act on unethical behavior requires a different mindset and triggers other cognitive processes than it is the case for ethical behavior. Here, the ability to self-justify behavior depends on the (un)ethicality of the (observed) act to the extent that it varies the boundaries of the moral wiggle room (Dana, et al., 2007). Empirical support is provided by the slippery-slope effect (cf. Gino & Bazerman (2009), Welsh et al. (2015)) as well as results on the "broken windows effect" (cf. Beckenkamp et al. (2014), Lefebvre, Pestieau, Riedl & Villeval (2015)), which suggest that unethical behavior is likely to be more contagious than unethical behavior.

We introduce $\boldsymbol{k}$, which represents a factor biasing behavioral adaptation towards the observation of unethical behavior. We assume that the degree of the adaptation bias depends on the difference between one's own and observed unethical behavior $\boldsymbol{\theta^i - \psi_t^i}$. With $\boldsymbol{\psi_t^i}$ still retaining its properties from **(2)**, the individual's maximization is of the form:

(7) $$\max_{x_t^i} U_t^i = I - \left(1 - \alpha_t\right) * \left(x_t^i - \theta^i\right)^2 - \alpha_t * k * \left(x_t^i - \psi_t^i\right)^2$$

with:

(8) $$k = \max(\theta^i - \psi_t^i + 1, 1)$$

This definition illustrates that whenever one's own initial behavior is more ethical than what is being observed from the peer behavioral contagion is stronger than observing more behavior that is more ethical. The strength of the difference in this contagion force depends on the

distance between own and observed behavior. This implies the following maximization of the individual's utility:

(9) $$\frac{\partial U_t^i}{\partial x_t^i} = 2\alpha_t \ \mathrm{k} * \left(\psi_t^i - x_t^i\right) - 2 * \left(\alpha_t \ - 1\right) * \left(\theta^i - x_t^i\right)$$

$$\Leftrightarrow x_t^i = \frac{\theta^i - \alpha_t \ \theta^i + \alpha_t \ \psi_t^i k}{\alpha_t \ k - \alpha_t \ + 1}$$

with the comparative static:

(10) $$\frac{\partial x_t^i}{\partial \alpha_t} = \frac{k*(\psi_t^i - \theta^i)}{\left(\alpha_t \ k - \alpha_t \ + 1\right)^2}$$

Under these assumptions, the resulting adaptation gap looks as follows:

(11) $$\left|x_2^i - \psi_2^i\right| = |\ldots| = \left|-\frac{(\theta^i - \psi_2^i)*(\alpha_2 \ - 1)}{\left(\alpha_2 \ k - \alpha_2 \ + 1\right)^2}\right|$$

We can easily see that the slightly differently appearing adaptation gap retains the same properties as in the symmetric model. In conclusion, we expect behavioral adaptation to be driven by social proximity with a bias towards unethical behavior. Put differently, we expect the spillover of unethical behavior to be more pronounced, thus leading to a stronger degeneration of behavior relatively to the rise of good Samaritans.

6. The Experiment

6.1 Experimental Design and Procedure

In order to study behavioral contagion, we are mainly interested in answering two questions with our experimental design: first, whether individuals revise their initial behavior in light of observing peer behavior. Second, whether and how a behavioral change depends on both the ethicality of the observed behavior and the social identification with the observed peers. Our design draws on a unique approach to study behavioral contagion in the lab that allows us to account for potential confounds potentially inherent into peer effects studies as argued by Manski (1993) and Angrist (2014) (see Thöni & Gächter (2015) for a similar approach). In order to be regarded as behavioral contagion, revised behavior has to be more similar to observed

behavior than one's initial behavior that one has decided upon prior to learning peer information. That is, revision of one's initial behavior must follow the direction of observed initial peer behavior.

Our basic design follows this straightforward procedure: action – observation of a peer – reaction.[34] Consider a variant of a two-player dictator game in which the participant (dictator) is matched with a charity (recipient). The dictator's action space entails taking away money from the charity, leaving the initial situation unchanged, or give money to a self-chosen charity (the basic design follows List (2007) and Bardsley (2008)). In following Eckel & Grossman (1996), we use a charity to increase the saliency of the involved decisions. The experiment is played one-shot with a possibility to revise one's initial behavior. Between the initial decision and potential revision, individuals are given the opportunity to observe the initial behavior of another random participant. Alongside the actual behavior, treatment variations include the alteration of unveiled social proximity information of the observed participant.[35] That is, in addition to learning actual behavior and the amount that was taken away or given by this participant, additional information on the participant's social proximity to oneself is varied with the random treatment assignment. The treatment variation lies in the information given about the social proximity to the observed peers: no information on proximity (Baseline), as well as high proximity (T1) and low proximity (T2) information. Proximity is calculated based on overlapping answers in the list of statements used in the beginning of each session and then presented to the participants in the form of below- or above-average proximity information to the observed peer.[36] We capitalize on a shortened 25 items list of statements compiled from a major US American dating website to ensure the validity of the questions in successfully matching people (see Gibbs, Ellison & Heino (2006) and Hitsch, Hortaçsu & Ariely

[34] Note that in order to exclude any hedging concerns throughout the whole experiment, information about the specifics of the design were only provided where necessary in order to reach a deliberate decision. That is, at Stage 1 participants were neither aware of the possibility to observe peers later on nor to revise their initial decision, ensuring unbiased initial behavior.

[35] Henceforth, we will use the terms social proximity and social identity interchangeably.

[36] The implementation of the low- and high-proximity information followed a very straightforward calculation. For each participant of the active group, an individual proximity score to both participants of the passive group was calculated based on overlapping answers in the list of statements. From each active participant's individual perspective, the passive participant with the higher (lower) score was labeled as the high (low) proximity peer. In fact, this calculation approach allows for the same passive person to be of high (low) proximity to one active person, while being of low (high) proximity to another active person, thus truly randomizing information. We abstained from providing explicit matching scores or percentages to retain maximum control. In addition, this allows us to alleviate the *false-consensus* effect, in which people systematically overestimate the degree of similarity to others. The provision of social cues of this kind allows the participants to update their beliefs reliably with respect to the actual degree of similarity. See Ellingsen & Johannesson (2008, p. 995) for a discussion.

(2010) for a discussion).[37] Since the business concept of dating websites is based on achieving high matching success rates, the use of such validated questions improves the success of incepting social identification between participants in the lab.[38]

The experimental procedure is represented by a single iteration of the following three stages:

First Stage - The Action: Starting with an equal distribution of money, each individual decides whether to (i) donate own money to the charity's account, (ii) not change the initial equal distribution, or (iii) take money from the charity and add to one's own account.

Second Stage – The Observation: Each active player observes one passive player of random who has engaged in either ethical or unethical behavior. In all three treatments, the exact information entails the monetary amount taken away from or given to the charity. Except for the baseline condition, observers received additional information related to the social identity to the observed peer, stating simply that based on the initial answers this peer is of higher or lower proximity as compared to the other passive peer that one cannot observe.

Third Stage – The Reaction: After observation, the active player is given the choice to revise his initial decision.

The experiment was concluded with a battery of non-incentivized questions to elicit attitude towards, among others, charitable giving and risk.

Our design accounts for the aforementioned reflection problem by randomly assigning all participants into two groups: *active* and *passive*. The group assignment is relevant to the action space available to the participants. If assigned to the *active* group, the participant is in the role of the *observer* and receives the opportunity to revise his initial decision. If assigned to the *passive* group, the participant is in the role of the *observed* person and is neither allowed to observe others nor to revise his own initial decision. In other words, in the *active* group, participants are free to request information about the behavior of one randomly chosen peer

[37] We should stress the fact that we report lower-bound results. We induce social proximity in a very simple way by providing participants with either the high- or low-proximity signal in the social identification treatments 1 and 2. Although this approach allows us to provide a comprehensive set of information to induce even more salient and distinct forms of social identity (i.e. by providing the exact matching score, the exact answers to the questions, letting participants put different weights on questions according to their individual importance and so on), we resort to this easy-to-use-easy-to-reproduce approach. See Appendix B for the exact list of questions used in our experiment.

[38] In addition, we elicit the strength of social identity to the observed peer using a variant of the self-evaluation scale of one's social identity following Luhtanen & Crocker (1992) to verify the robustness of our social identity implementation approach. The non-parametric and regression estimations yield the same overall results both in direction and in magnitude.

following their initial decision and are then given the chance to revise. Whereas in the *passive* group, the participant's initial behavior is held fixed and no further decisions are made.

Payoff structure: Importantly, to exclude any form of strategic interaction that might potentially dilute results or affect the saliency, the participant's decisions only affected one's own and the chosen charity's payments but not those of other participants. That is, each individual's decision had no monetary impact on other individuals, therefore a change in behavior was due purely to behavioral contagion and not due to other-regarding concerns.[39] This becomes even more salient by randomly picking one of the individuals at the end for which the behavioral decision was implemented, while everyone else received a flat income irrespective of his actual behavior. In monetary terms, each participant and the respective charity received the ECU equivalent of 15 Euro, thus allowing a participant to leave with a maximum (minimum) of 30 (0) Euro if the participant decided to take away all the money from (give all the money to) the charity. In order to increase the saliency of ethical behavior, we added a multiplier to the setting. That is, the experimenter doubled all Euro remaining in the charity's account at the end of the experiment. After all decisions have taken place, one participant was chosen at random and the respective decision was implemented with respect to taking from or giving to the charity, while every other participant in the session received a flat income of 7.50 Euro including a show-up fee of 2.50 Euro.

Several points are worth noting. In order to retain maximum control and reduce heterogeneity in observed behavior, in each session exactly two participants were chosen at random as passive (that is, being observed by peers), while participants randomly chosen as being active (that is, observing peers) always observe only one of these two passive players. What is more, the treatment differences are based solely on the social proximity information. Importantly, the observing peers received the information that a random draw will determine whether they will observe a high or low proximity peer and that this person would have either taken money away from or given money to the charity. The observer would then randomly learn the behavior of exactly one passive player but was never able to infer other participants' behavior from this information, neither active nor passive.[40] This double-random procedure ensured

[39] To some degree, the experimental design resembles the theoretical considerations of Alós-Ferrer & Schlag (2009).

[40] Prior to the actual observation, we elicited incentivized beliefs about the behavior of the two passive participants. Those beliefs ended up being irrelevant in predicting the active participants' behavior. We will return to this point shortly in the analysis section.

that each active player's information set would be restricted to the peer's observed behavior and focus the participant's attention on processing this information. In addition, to allow for perfect comparability across treatments, participants were randomly assigned to one of the three treatments within the same experimental session. In sum, this design allows us to measure clean peer effects every time a participant revises his/her initial decision after observing peer behavior.[41] The design is illustrated below:

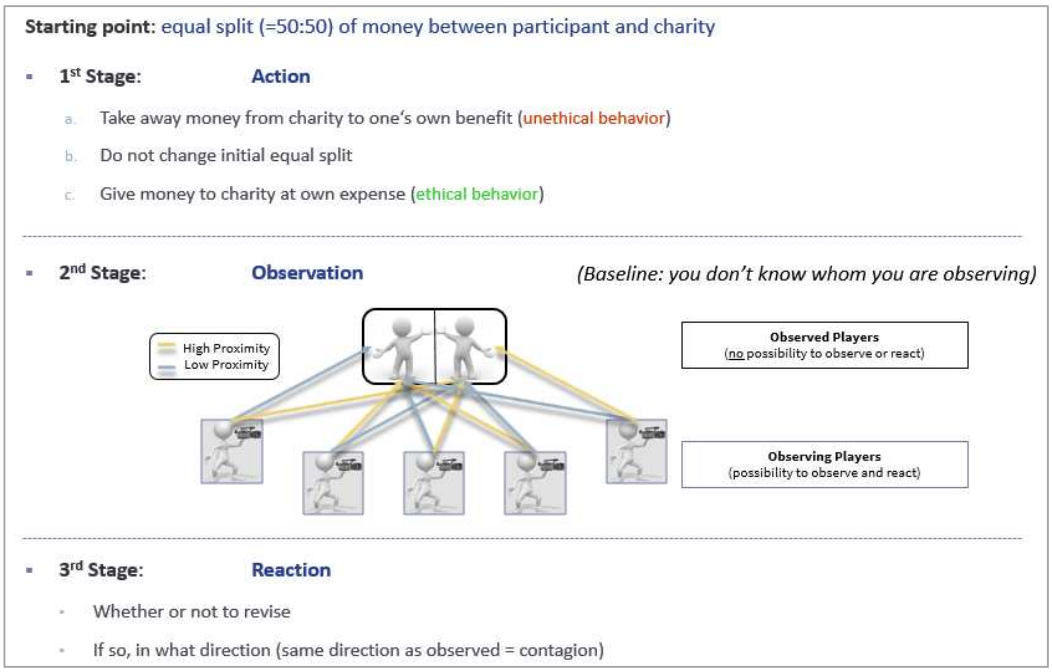


Figure 1: *Experimental Design and Procedure*

[41] With reference to the different concepts in economics and (social) psychology explaining behavioral contagion (as discussed in chapter 4) and in light of our design, some theories are better at explaining behavior in our experiment than others are. Contagion observed in our experiment is likely to be explained by the theories of *social decisions and social distance* (Akerlof (1997), Glaeser & Scheinkman (2004)), *imitation of preferences* (Sliwka, 2007), as well as by some of the theories in (social) psychology, such as *social learning* (Bandura, 1971), *norms* (Cialdini et al. (1990) and Bicchieri (2006)) and *self-expansion* (Aron & Aron, 1986). Other theories are not applicable due to reasons of absence of learning better outcomes (which is required by Alós-Ferrer & Schlag's (2009) imitation concept), or the ability to not only observe but also be observed by the peers (which is required by, among others, Bernheim's (1994) theory on taste for conformity). While it is questionable whether it is possible at all, we shall not attempt to disentangle clearly, which theory best explains behavioral contagion observed in our experiment.

6.2 Hypotheses

In order to generate hypotheses that are sound with existing theory, we derive our theoretical model as it is discussed in chapter 5 and our predictions from previous research.[42] The slippery-slope argument (cf. Gino & Bazerman (2009), Welsh et al. (2015)) as well as results on the "broken windows effect" (cf. Beckenkamp et al. (2014), Lefebvre, Pestieau, Riedl & Villeval (2015)) suggest that unethical behavior is likely to be more contagious than unethical behavior. We thus derive our first hypothesis as follows:

H_1: ***Unethical behavior is more contagious than ethical behavior.***

Previous research indicates that social identification is a decisive predictor of behavior in different contexts related to charitable giving, trust, punishment, and reciprocity (cf. Leider et al. (2009), Chen & Li (2009)). What is more, some of the existing research on neighborhood effects (cf. Damm & Dustmann (2014)) and our theoretical predictions from chapter 4.2 support these assumptions. Consequently, we formulate our second hypothesis:

H_2: ***Social identification amplifies the contagion of behavior in general and unethical behavior in particular.***

We are also concerned with understanding the main driver of behavioral contagion by answering whether the unethicality of observed behavior or social identification to one's peers is a stronger predictor of behavioral contagion. While different streams of research suggest the relevance of both channels individually, we are the first, to the best of our knowledge, to compare these channels in terms of their impact on the magnitude of behavioral contagion. We follow the existing research on social identity (cf. Tajfel (1982), Hoffman, McCabe & Smith (1996), Bohnet & Frey (1999), Eckel & Petrie (2011), Charness & Gneezy (2008)) and expect behavioral contagion to be more pronounced the higher the social identification with one's peer is, independent of the (un)ethicality of the observed behavior. We thus derive our final hypothesis:

H_3: **Social identification is the main driver of behavioral contagion.**

[42] Our hypothesis will mainly focus on two behavioral traits that we can observe in our experiment: the first is individual $\boldsymbol{i}$'s revised amount after observing a peer's initial decision: $\boldsymbol{x_2^i}$. The second is the adaptation gap of individual $\boldsymbol{i}$ and $\boldsymbol{j}$ after $\boldsymbol{i}$'s revision decision: $\left|\boldsymbol{x_2^i} - \boldsymbol{\psi_2^i}\right|$. We will capitalize on these two aspects to support our hypotheses in the results section.

In what follows, we will present and discuss the results to our hypotheses 1-3.

6.3 Results and Discussion

We conducted the experiment at the BaER-Lab at the University of Paderborn, Germany. Participants were recruited using ORSEE (2004). We used zTree (Fischbacher, 2007) to run our experiment. In sum, 227 participants throughout 9 sessions were randomly assigned to one of the three main treatments (no proximity, high proximity, and low proximity) as well as to one of the two sub-treatments (observing either ethical or unethical behavior).[43] Each session lasted about 45 minutes and the hourly average earnings were €10.50. What is more, an average of €30 was donated per session to various charities.[44]

Figure 1 details our behavior data in two ways: before observation in pooled form and after observation by treatment. Starting with a descriptive observation, we observe some heterogeneity across treatments in terms of the distribution of (un)ethical behavior. On the x-axis, we depict the amount of money that was allocated to the charity. The point of departure is 300, which represents the equal *a priory* distribution of money between the participant (300) and the charity (300). Thus, any value below (above) 300 depicts the individual taking from (giving to) the charity's account.

Prior to observing peer's behavior, decisions are mainly clustered at 300 ECU, which represents the decision to not change the initial equal distribution between oneself and the charity. Across all treatments, a total of 25% of participants decided to revise his/her initial decision. After the observation, however, we find a perceptible skewness towards unethical behavior, particularly in the high proximity condition. Although observing either ethical or unethical behavior is equally likely, this finding provides us with a first indication that taking (unethical behavior) is more contagious than giving (ethical behavior). We will return to this argument shortly in our hypothesis $\mathbf{H_1}$.

[43] Out of the 227 participants, 18 (8%) participants were randomly assigned to being observed and thus remained passive after their initial decision, and 24 (11%) decided to opt-out and not to observe peer behavior. The latter represents a significant portion of participants that refuse to learn about peer behavior and thus speak to the aforementioned problem of forcing participants to observe peers. See chapter 5.1 for discussion. In total, this leaves us with 185 observations entering our analysis.

[44] See Appendix A for more details.

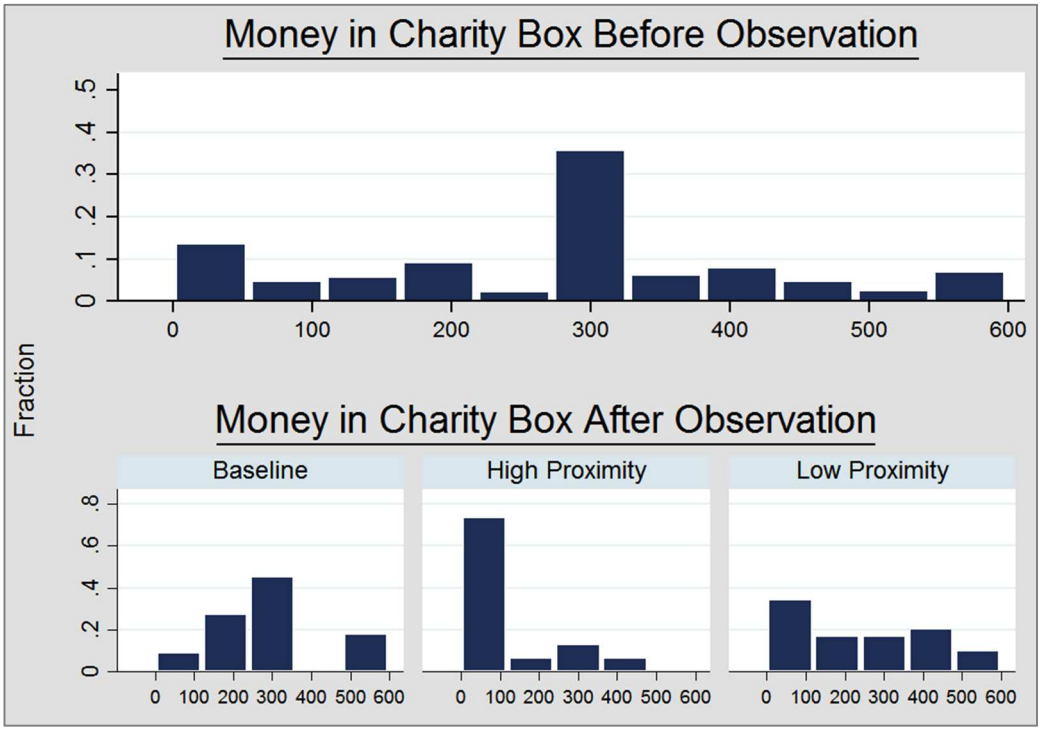


Figure 2: *Incidence of choices to not change the initial equal distribution / give money to / take money from the charity before and after peer observation. The horizontal axis depicts a continuum of ECU left in the respective charity's cash account, with 300 representing the starting allocation. The choices were clustered in the figure for the sake of readability. The vertical axis depicts the fraction of participants indulging in the particular behavior.*

Next, we turn to testing our hypotheses.[45] We are interested in whether unethical behavior is more contagious than ethical behavior. We will do so by illustrating the amount revised (%) after observation purely depending on the (un)ethicality of observed behavior.[46] More specifically, for the *amount revised (%)*, the results suggests that the (un)ethicality of observed behavior indeed affects the individual's revision decision. When observed ethical (unethical) behavior, meaning that participants observed a peer who gave money to (took money from) the charity, participants gave on average 6.42% more (17.3% less) money to the charity as compared to their initial behavior prior to having observed a peer's behavior. The results are significant at the 1% level (p = 0.000, z = 4.365) and indicate a change in behavior almost three

[45] In general, non-parametric Mann-Whitney-U (MWU) specifications are used throughout this paper to test for differences among those who decided to revise initial behavior. Unless noted otherwise, these findings are robust to alternative non-parametric specifications, such as the Kolmogorov-Smirnov test, and varying examinations of missing observations.

[46] For *amount revised (%), a* positive (negative) value implies that individual $\boldsymbol{i}$ has given more money to (has taken more money from) the charity as compared to his/her initial behavior prior to observing a peer's behavior. Robustness checks involve the analysis of differences in the *adaptation gap (%).* For adaptation gap (%), a value below (above) 1 implies that the adaptation gap, denoted as $\left|x_2^i - \psi_2^{ij}\right|$, has gotten smaller (larger) after individual $\boldsymbol{i}$ had the chance to revise his/her initial decision. A value of exactly 1 means that the adaptation gap remained the same after the revision stage as compared to the initial decision. This could be either due to individual $\boldsymbol{i}$ deciding not to revise his/her initial decision or due to a revision which is equidistant in monetary terms. The results support $\mathbf{H_1}$ and are presented in Appendix A.

times as large when unethical behavior was observed as compared to ethical behavior. The results are highly suggestive of unethical behavior being more contagious. Thus, these results strongly support our hypothesis **H_1** and confirm that unethical behavior is indeed more contagious than ethical behavior independent of social identification.[47] Figure 3 illustrates our findings.

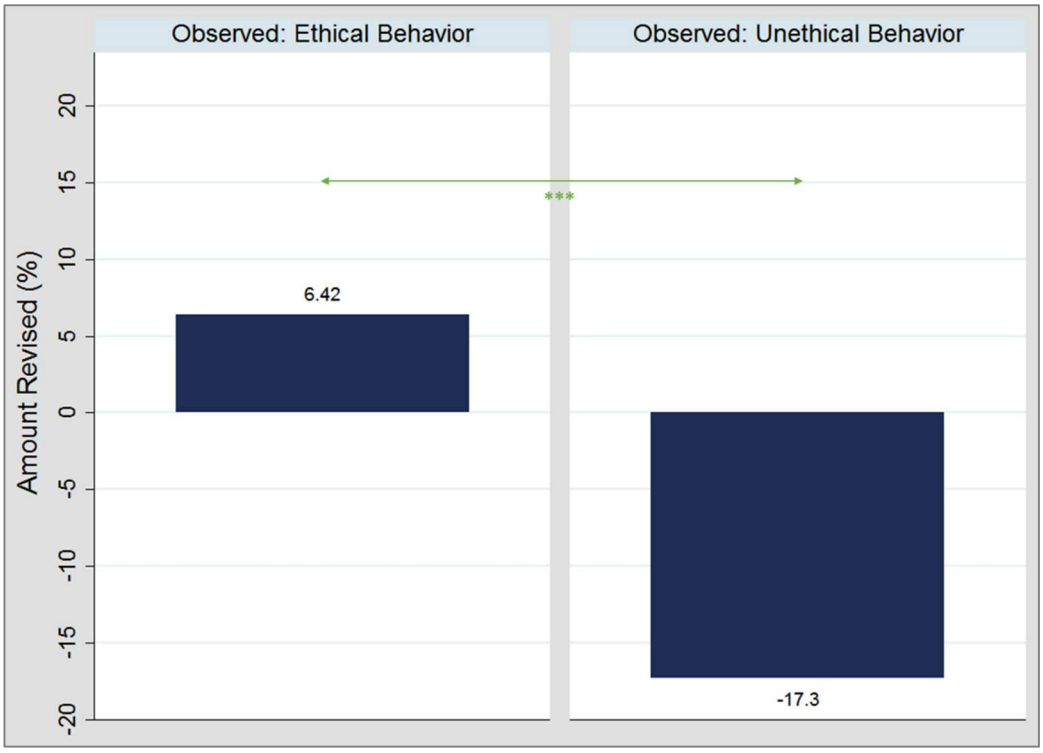


Figure 3: *Amount Revised (%) and Observed (Un)Ethicality. The figure depicts the amount revised as percentage of one's initial behavior. Any value above (below) zero indicates that more (less) ECU were given after the revision to the charity relative to one's initial decision. The analysis is broken down into the (un)ethicality of observed behavior by the active participants.*

We are also interested in the type of behavioral contagion that is triggered by social identification. Following hypothesis **H_2**, we assume that social identification amplifies the contagion of unethical behavior in an over-proportional way as compared to contagion of ethical behavior. We thus examine the role of social identification in affecting the magnitude and direction of revision choices. Both findings are illustrated below in figure 4.

Our results robustly indicate that higher social identity indeed triggers stronger behavioral contagion, particularly contagion of unethical behavior. As social identity increases, the mag-

[47] Unless noted otherwise in the results section, we obtain similar results in terms of evidence and significance when using absolute ECU numbers instead of percentages.

nitude of revised behavior increases as well, peaking at -29.8% for the high proximity condition. The differences in behavior are significant at the 1% level (p = 0.000, z = 4.759) when comparing behavior in the unknown proximity with the high proximity condition. Likewise, the results are significant at the 1% level (p = 0.000, z = -3.448) when comparing the high proximity with the low proximity condition. Here, again, the negative numbers of amount revised suggest that, in terms of magnitude of revised behavior, unethical behavior is strongly pronounced and thus more contagious than ethical behavior. Overall, we find ample support for our hypothesis **H_2** and thus conclude that the magnitude of a revision of one's initial behavior is indeed strongly correlated with social identification.

Consistent with our theoretical model, the results also yield strong support for the idea that the reduction in the adaptation gap is driven by social identification. As predicted by α^{ij}, the stronger the social identification to one's peer is a robust predictor (all at the 1% level) of one's adaptation gap to the observed peer after observation. By the numbers, we obtain p = 0.000 and z = 6.104 (p = 0.000 and z = -3.441) when comparing unknown proximity versus high proximity condition (high proximity versus low proximity condition). We find that the adaptation gap is inversely correlated to the social identity.

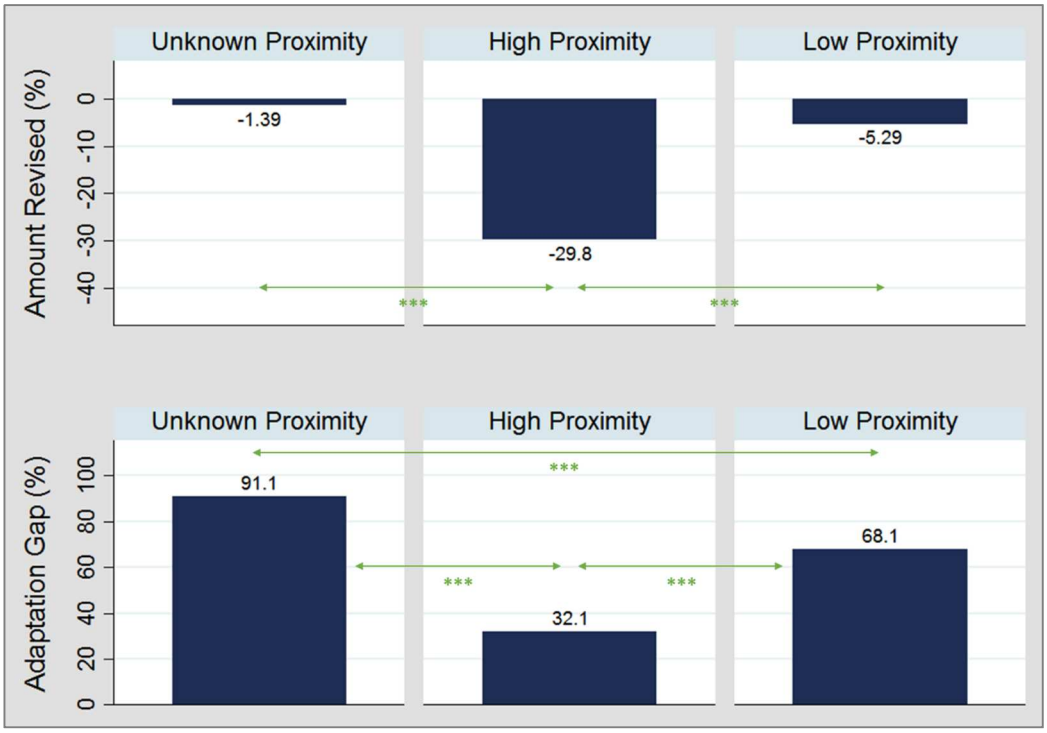


Figure 4: *Amount revised (%) and adaptation gap (%) by social proximity.*

Along similar lines, behavioral contagion as a function of both social identification and the (un)ethicality of observed behavior is illustrated in figure 5. The behavioral space is described by three alternatives and also speaks to the (non-)existence of observed peer effects:

- **No Contagion:** after observing peer behavior, the participant either did *not* revise his/her initial behavior or revised it into the opposite direction of what he/she observed the peer has done.
- **Contagion:** after observing peer's behavior, the participant did revise his/her initial decision. The revision was directed into the direction of the observed behavior. This behavior indicates the existence of behavioral contagion caused by peer effects.

The breakdown of behavioral changes by different levels of social proximity provides additional evidence for hypothesis **H_2**. When comparing to the condition where no social proximity to one's observed peer was induced, behavioral is more contagious for both low and high proximity situations in both the ethical and unethical domain. Again, since we report lower-bound results, this is a strong indication that our method of inducing social identification works and is likely to produce even stronger results when social proximity would be introduced in a more sophisticated way.

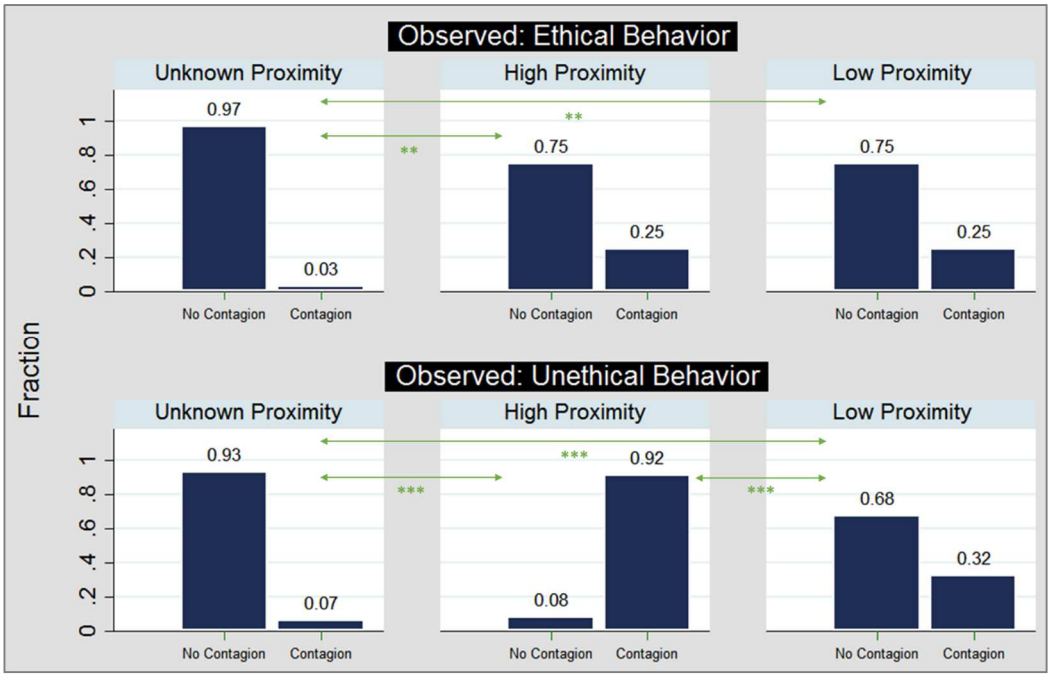


Figure 5: *Behavioral change by treatment and observed (un)ethicality. This figure illustrates the fraction of participants exhibiting behavioral contagion broken down into both the observed (un)ethicality of behavior and the social identity to one's observed peer. Unless the active participant revises his/her behavior into the direction of observed behavior the behavior is not classified as contagion.*

In terms of explaining the drivers of behavioral contagion, two literature streams have not yet been brought together in existing research: the role of social identification on behavioral contagion on the one hand, and the role of observed behavioral unethicality on the other. Our experimental design allows us to do exactly this for the first time and ascertain the main driver of behavioral contagion. Following our previous discussion, we assume social identification to be a stronger driver of behavioral contagion than the unethicality of observed behavior (**H_3**). As a first, we investigate whether behavioral contagion is different under varied levels of social proximity when directly comparing contagion in the ethical versus the unethical domain using non-parametric comparisons. Our results provide strong support that behavioral contagion is asymmetric. In particular, a variation in social identification yields no significant variation in behavioral changes in the ethical domain. However, the results are strongly statistically significant when looking at behavioral changes as a function of social identification in the unethical domain. Here, when comparing the high proximity condition to the no proximity (low proximity) condition, the Mann-Whitney-U statistics yield results that are highly significant at the 1% level with p = 0.000 and z = 6.025 (p = 0.000 and z = -4.005). These results highlight the importance of our contribution in this paper: peer effects are not uniformly in place, but rather strongly depend on both the (un)ethicality of observed behavior and the degree of social identification to the observed peer.

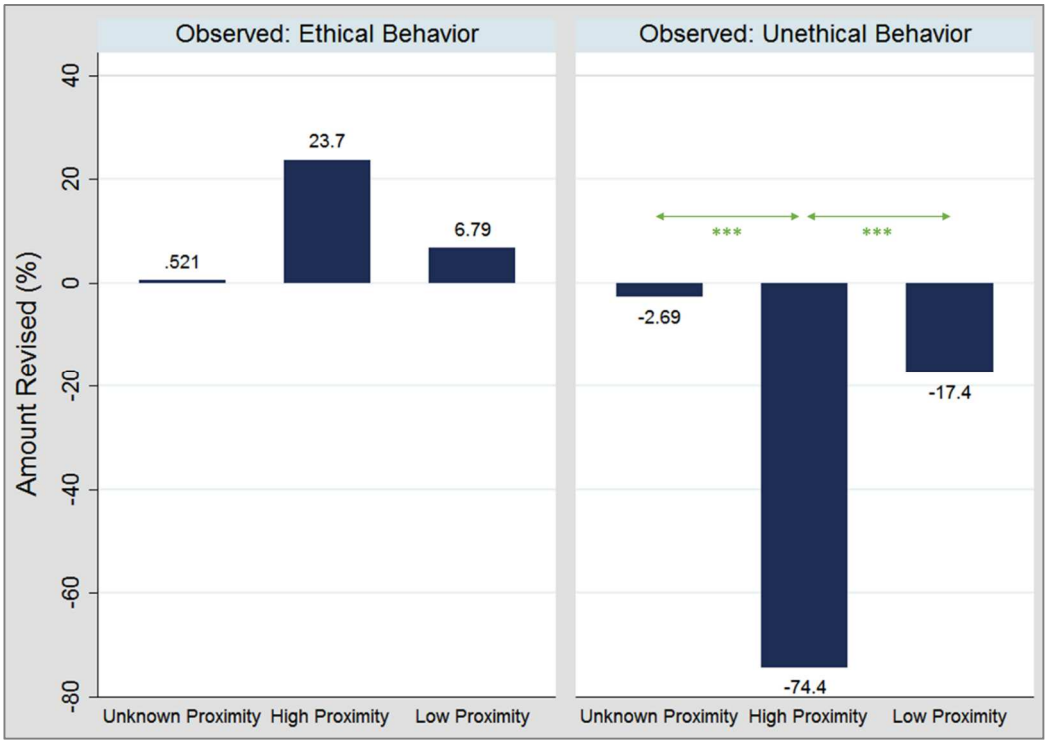


Figure 6: *Amount revised (%) by social proximity and observed (un)ethicality.*

What is more, we ran several regressions, including OLS, Logit, and Tobit, where applicable, in order to assess the robustness of our results. In sum, across different specifications our results strongly suggest that the observation of (un)ethical behavior does not trigger any particular behavior, neither ethical nor unethical. Thus, the mere observation of behavior alone is insufficient for the existence of peer effects in the (un)ethical sphere, but is rather dependent on the social identification to one's peers. These findings are in support of our hypothesis **H_3**. Findings are presented in Appendix A.

All in all, we find convincing support for all of our three hypotheses: unethical behavior is indeed more contagious than ethical behavior (H_1), social identification drives the magnitude of behavioral contagion (H_2), and social identification is a more reliable predictor of behavioral contagion than the observed (un)ethicality of observed behavior (H_3). Here, it is important to stress a particular point with respect to the interpretation of the results. Arguably, the experiment's framed environment created by the presence of a charity could potentially carry a norm of prosocial behavior in the lab in and of itself. That is, prior to observing one's peers, some participants at the margin of behaving unethically might carry the normative belief that taking from a charity represents inappropriate behavior and thus initially refrain for it. If so, it would come as no surprise to observe stronger contagion of unethical as compared to ethical behavior because those who wanted to behave unethically in the first place but refrained from doing so might now find justification in their peer's behavior. In this respect, two important remarks should be made: first, if anything, such an assumption would only explain level effects but not treatment differences because such beliefs are by experimental design uncorrelated with the treatments. Consequently, irrespective of the existence of potential norms, our design renders our main finding valid: behavioral contagion is highest where social identity is strongest.[48] In addition, we elicited incentivized beliefs about what participants thought about his peer's behavior prior to observing it. As the regressions results suggest, such beliefs yielded no explanatory power and thus play no role in neither explaining the magnitude nor the differences of behavioral contagion.

[48] I would like to thank Gary Bolton, René Fahr, and Elena Katok for point this out and for related discussions.

From a policy perspective, our results stress that social proximity renders it difficult to change individual behavior, but it rather amplifies one's initial (un)ethicality. We will return to this point in our policy recommendations in chapter 7. In light of the very conservative inception of social identification of providing very limited information on social identification, we deem these results to represent a lower bound thus strengthening the role of social identification within the context of behavioral contagion. Our lower-bound approach comes from the fact that participants were neither told the exact matching percentage nor the actual interests and preferences they had in common. If having been randomly assigned to one of the social identification treatments, participants only knew whether they were observing a peer with above or below average congruence. A more sophisticated way to induce social identification and match participants accordingly is likely to produce results that are more pronounced. The well-engineered mechanisms implemented by dating websites to match people and achieve high success rates are a shining example for what is possible: excluding matching partners based on personality traits that represent a no-go (e.g. smoking), putting emphasis on particular interests (e.g. sports, religion), or individual characteristics (e.g. looks, education). We deliberatively refrained from applying sophisticated measures of this kind and rather resorted to an easy-to-use-easy-to-reproduce methodological approach that could be used in future experiments in which inducing salient social identity is key.

Along these lines, the presence of a potential experimenter demand effects (EDE) is worth addressing since their presence has potentially been problematic to prior peer effect studies (for a discussion see Thöni & Gächter (2015)). Because we are mainly interested in treatment differences rather than in overall levels, the experimenter demand effect is deemed less problematic as long as its existence and magnitude is orthogonal to the treatment variation (Zizzo, 2010). Nonetheless, we considered existing experimental studies to rule out experimenter demand effects to the extent possible. Not exclusively to but prominently existing in peer effect studies, forced learning (i.e., forced observation of one's peer's behavior) might potentially induce EDE or even lead to resentment on the side of the participants. Forced observation might trigger thoughts related to being expected to use the information to reconsider and potentially revise initial behavior. In previous general and peer effect studies in particular, this issue has normally been overlooked, mainly to avoid self-selection problems. However, when the option not to learn is withheld from participants, the obtained results are potentially confounded. We deem this challenging to the study of peer effects and should thus be discussed.

Existing research indicates that individuals sometimes choose to deliberately remain ignorant about the state of nature (see Carrillo & Mariotti (2000), Dana, Weber & Kuang (2007), Conrads & Irlenbusch (2013), Bartling, Engl & Weber (2014), and Grossman (2014)). If present, such strategic ignorance might be an important component of our experiment. A potential reason not to acquire costless information is related to the avoidance to indulge in negative self-image updating or guilt aversion (cf. Wells (1978), Baumeister (1998), Charness & Dufwenberg (2006)). One can plausibly assume that such aversion is even stronger when studying peer effects within an (un)ethical dimension. Thus, forcing participants to learn potentially unpleasant information might lead to biased behavior and might even increase EDE. In order to address this challenge, our design follows Conrads & Irlenbusch (2013) and Bartling, Engl & Weber (2014) and draws on a mechanism in which learning peer information is voluntary (see also Eckel & Petrie (2011)).[49] Additionally, in order to rule out any reputational concerns, social learning, or reciprocity, the experiment includes an anonymously played one-shot interaction with another participant. Such an experimental design allows us to study behavioral contagion in the lab in an unbiased way. To our knowledge, this represents a novel design in studying peer effects in the lab in general and the behavioral contagion of (un)ethical behavior in particular while controlling for potential confounds that peer effect studies suffer from regularly (see Manski (2000), Falk & Fischbacher (2002), Angrist (2014)). In the light of our experimental set-up, any treatment-specific information is provided only after one's deliberate decision to learn peer behavior. Thus, any still potentially existing form of EDE would be fully uncorrelated with the treatments and thus exhibit only a general level-effect, if any.

7. Lessons Learned: Policy Implications

As argued before, understanding social interactions in general and in particular the potentially resulting peer effects is fundamental from a policy perspective. It does not only help to understand societal and economic outcomes beyond what standard economic forces can explain (i.e., the massive surge in female labor participation rates in World War II (Mulligan, 1998) or the escalation of crime rates (Levitt, 1999)). It also allows us to implement better-targeted policy measures to tackle a battery of challenges such as reducing crime rates, improving

[49] However, deliberately allowing participants to remain ignorant about peer behavior bears the risk of self-selection effects. It is worth noting that in our experiment 10% of all the participants decided not to acquire peer information. Importantly, however, this choice is unconditional on the participant's initial behavior, thus strongly suggesting the absence of any self-selection mechanism leading to potential biases in our analysis.

health conditions, or increasing labor market participation. “To the extent that theory and measurement of social interactions enables us to understand these massive changes, the study of social interactions potentially has major policy relevance” (Glaeser & Scheinkman, 2004, p. 84). In this chapter, we will discuss some of the policy implications that one can infer from our results. We will follow the main theme of this paper and approach this topic from two sides: the ethical and unethical context (for a broader discussion see Irlenbusch & Villeval (2015)).

Starting with the ethical perspective, voluntary redistribution of income e.g. in the form of charitable giving is an integral part of humaneness, with up to 90% of Americans donating to charities. Understanding the drivers of charitable giving, such as altruism and social pressure, has been at the heart of last decade’s research, both in the field and in the lab (cf. Levitt & List (2009), DellaVigna, List & Malmendier (2012)). Still, we seem to have an imperfect understanding of what really motivates giving (Andreoni, 2006). Beyond what we have since learned about the existence of peer effects with respect to contribution decisions (see discussion in chapter 2 and 3), our experiment yields several new insights helpful to understanding the extent to which such behavior shows up, especially as compared to unethical behavior. We will return to this comparison shortly.

In turn, unethical behavior in its various forms impairs the daily life. Exemplarily, yearly global tax evasion ranges at an abstruse $3.1 trillion or 5.1% of world GDP (The New York Times, 2011), over $1 trillion is estimated in bribes paid yearly around the globe (The World Bank, 2013), and some 210 million people use illicit drugs each year (United Nations Office on Drugs and Crime, 2011). Here, one might credibly argue that from a purely rational self-maximizing perspective we should observe way less illicit behavior than we actually do. From a game theoretic perspective, in some illicit deals that involve trust-related actions such as bribery do not represent a subgame perfect Nash equilibrium, which is true for both one-shot and finitely repeated contexts. Thus, in many situations the occurrence of illicit behavior is already surprising (for a discussion, see Dimant & Schulte (forthcoming)). A solution to this conflict is the recognition of, among others, peer effects. As has been thoroughly argued throughout the paper, the incorporation of behavioral contagion allows us so explain why observed behavior goes seemingly beyond clear-cut self-maximization, but is rather embedded in and the result of a social context.

Consequently, both aspects raise the following question: how do these findings translate into a real world setting and what to do about it? While being careful at drawing concrete inferences from a laboratory setting and relating them directly to the outside world, our results conclusively indicate the spillover of unethical behavior to be much more likely than the spillover of ethical behavior. That is, getting people to start donating solely based on peer effects (e.g. through observing others giving to charity) has a long way to go compared to having them do something unethical. What we see, however, is the individual's responsiveness to social identification, in particular on the side of females. A potential solution to have people donate more to a good cause is to provide them information beyond simple statistics on what other people do, i.e. amount of money that has been collected so far (like Wikipedia). Instead, some research already indicates that the inception of social norms especially for settings that most closely match the individual's immediate situational circumstances have the strongest effect on compliance (Goldstein, et al., 2008). . Our results suggest to go one step further and provide information conveying social proximity to the peers, thus increasing the saliency of social identification. Exemplarily, a message of the form "People in your neighborhood / with similar demographic characteristics have donated an average of $..." would lead to pick-up rates of this behavior that are higher than when resorting to a simple statistic.

In all likelihood, a similar approach could be applied to make people refrain from behaving unethically. Research on slippery slope indicates that once behavior is spoiled, even honest people converge quickly to a steady state with a plethora of unethical behaviors (cf. Gino & Bazerman (2009), Welsh et al. (2015)). Even more worrisome, recidivism rates for convicts are normally very high, leading to what is called a recidivism nightmare. In the US recidivism rates are up to 80% of re-arrests within the first 3 to 5 years after their release from prison (National Institute of Justice, 2014). This is particularly detrimental from a welfare perspective, as the US spends approximately $75 billion on incarceration and $260 billion on prevention, detection, and prosecution of crime on a yearly basis (Khadjavi, 2015). In terms of effective countermeasures to unethical behavior, our experimental results indicate that exposing individuals to unfaithful of socially close people is likely to trigger repulsion and less unethical behavior, thus contributing to a positive transformation.[50]

[50] For an approach along similar lines, see Pennsylvania State University's Justice Center for Research on desistance from crime (Pennsylvania State University's Justice Center for Research, 2014).

8. Conclusion and Outlook

Deviant behavior that benefits oneself at the expense of others is socially harmful and brings about second-best solutions that are distortive from a welfare perspective. Conversely, the voluntary redistribution of money to those who have the least, e.g. in the form of donations, is socially desired. It is worthwhile to understand the underlying mechanism that drives (un)ethical behavior in order to implement effective policy measures that mitigate or facilitate either behavior. Beyond pure self-maximizing considerations, behavior is also be the result of social interactions in which the conformity to particular behavior is affected by one's peers (Akerlof (1997), Glaeser & Scheinkman (2004)). One particular mechanism through which social interactions occur are peer effects, which play a decisive role in explaining societal and economic outcomes. A battery of behavioral traits affects the shape and magnitude in which social interactions occur. An extensive stream of literature suggests that individuals are social animals and care for esteem, respect, reputation, among other things (cf. Rabin (1993), Fehr & Schmidt (1999), Akerlof & Kranton (2000), Bolton & Ockenfels (2000), Charness & Rabin (2002), Bénabou & Tirole (2011)). As such, individuals steadily act and react in social environments that define their role and standing within the social group.

Although peer effects have been extensively studied in different contexts both in the field and in the lab (see discussion in chapter 3.3), research is still at the outset of understanding the role of peer effects in (un)ethical settings. It is this paper's goal to improve our understanding of whether, to which extent, and through which channels individuals are influenced by their peers to engage in (un)ethical behavior, in which they most likely would have not engaged otherwise.

This paper is in line with and an extension of the seminal Moving-to-Opportunity field research that has allowed us to understand the positive societal spillovers of having the chance to move to a better neighborhood. In going beyond, we try to understand the spillovers of both unethical and ethical behavior and the role of social identity to one's peers by employing a controlled lab experiment. Understanding behavioral contagion is important from a policy perspective in order to set the right incentives and effective measures to improve pro-sociality and mitigate unethical behavior.

By extending existing research from both the methodological and the content perspective, our work contributes to a better understanding of the nature of peer effects and behavioral

spillovers, and answers the following two questions in particular: for one, to what extent does the (un)ethicality of a peer's observed behavior influence one's own behavior? For another, what is the role of social identity to one's peers in affecting behavioral contagion? For our purposes, we extend a variant of the give-or-take dictator game as introduced by List (2007) and used by Bardsley (2008) by the use of an ethical setting. To provide clean evidence on peer effects, we capitalize on a one-shot dictator game in which participants are given the opportunity to give to or take money away from the charity before and after learning the (un)ethical magnitude of peer behavior. Treatment variations include different levels of social identity to the observed peers, which we incept by the novel use of a matching algorithm based on a series of dating website questions. We deliberatively refrained from applying sophisticated measures of this kind and rather resorted to an easy-to-use-easy-to-reproduce methodological approach that could be used in future experiments in which inducing salient social identity is key.

Our results suggest that a) unethical behavior is more contagious, and b) social identification with one's peers and not the (un)ethicality of observed behavior is the main driver of behavioral contagion. Our results conclusively yield that the mere observation of behavior alone is insufficient for the existence of peer effects in the (un)ethical sphere, but is rather dependent on the social identification to one's peers.

Beyond these first results, much more scientific research has to be done in order to generate reliable measures to achieve both, more ethical and less unethical behavior. Exemplarily, recent MTO-research points at gender differences in behavioral assimilation (cf. Chetty, Hendren & Katz (2015)). It is important to understand to which extent these differences are driven by the two factors studied in this paper: the (un)ethicality of peer behavior one is exposed to and the magnitude of social identity to one's peers. Our research is hopefully one of many more contributions to come.

9. Bibliography

[1] Akerlof, G., 1980. A Theory of Social Custom, of Which Unemployment May Be One Consequence. *The Quarterly Journal of Economics*, Volume 94, pp. 749-775.

[2] Akerlof, G., 1997. Social Distance and Social Decisions. *Econometrica*, 65(5), pp. 1005-1027.

[3] Akerlof, G. & Kranton, R., 2000. Economics and Identity. *Quarterly Journal of Economics*, 115(3), pp. 715-753.

[4] Alos-Ferrer, C. & Schlag, K., 2009. Imitation and Learning. In: P. Anand, P. Pattanaik & C. Puppe, eds. *The Handbook of Rational and Social Choice.* s.l.:Oxford University Press.

[5] Anderson, L. & Holt, C., 1997. Information Cascades in the Laboratory. *American Economic Review*, 87(5), pp. 847-862.

[6] Andreoni, J., 1989. Giving with impure altruism: applications to charity and Ricardian equivalence. *Journal of Political Economy*, Volume 97, pp. 1447-1458.

[7] Andreoni, J., 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, Volume 100, p. 464–477.

[8] Andreoni, J., 1995. Cooperation in Public-Goods Experiments: Kindness or Confusion?. *American Economics Review*, 85(4), pp. 891-904.

[9] Andreoni, J., 2006. Philanthropy. In: S. Kolm & J. Ythier, eds. *Handbook of Giving, Altruism, and Reciprocity.* (Amsterdam: Elsevier/North-Holland, p. 1201–1265.

[10] Andreoni, J. & Bernheim, D., 2009. Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica*, 77(5), pp. 1607-1636.

[11] Angrist, J. D., 2014. The perils of peer effects. *Labour Economics*, Volume 30, pp. 98-108.

[12] Angrist, J. D. & Lang, K., 2004. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5), pp. 1613-1634.

[13] Apesteguia, J., Huck, S. & Oechssler, J., 2007. Imitation - theory and experimental evidence. *Journal of Economic Theory*, Volume 136, pp. 217-235.

[14] Arcidiacono, P. & Nicolson, S., 2005. Peer Effects in Medical Schools. *Journal of Public Economics*, 89(2-3), pp. 327-350.

[15] Aron, A. & Aron, E. N., 1986. *Love as expansion of the self: Understanding attraction and satisfaction.* New York: Hemisphere.

[16] Aron, A., Aron, E. & Smollan, D., 1992. Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, Volume 63, pp. 596-612.

[17] Aronson, E., 2011. *The Social Animal.* 11 ed. s.l.:Worth Publishers.

[18] Asch, S., 1951. Effects of Group Pressure Upon the Modification and Distortion of Judgments. In: H. Guetzkow, ed. *Groups, Leadership, and Men – Research in Human Relations.* Pitsburgh: Carnegie Press, pp. 177-190.

[19] Asch, S., 1958. Effects of Group Pressure on the Modification and Distortion. In: E. Maccoby, T. Newcomb & T. Hartley, eds. *Readings in Social Psychology.* New York: Holt, Rinehart & Winston.

[20] Ash, S. E., 1952. *Social Psychology.* Englewood Cliffs: Prentice-Hall.

[21] Azmat, G. & Iriberri, N., 2010. The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment Using High School Students. *Journal of Public Economics*, Volume 94, pp. 435-452.

[22] Baldwin, J. M., 1894. Imitation: A chapter in the natural history of consciousness. *Mind*, Volume 3, pp. 25-55.

[23] Bandiera, O. & Rasul, I., 2006. Social Networks and Technology Adoption in Northern Mozambique. *The Economic Journal*, 116(514), pp. 869-902.

[24] Bandura, A., 1971. *Social Learning Theory.* New York: General Learning Press.

[25] Bandura, A., 1986. *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

[26] Banerjee, A., 1992. A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, 107(3), p. 797–817.

[27] Banerjee, A., Chandrasekhar, A., Duflo, E. & Jackson, M., 2013. The Diffusion of Microfinance. *Science*, 341(6144), pp. 363-372.

[28] Bardsley, N., 2008. Dictator game giving: altruism or artefact. *Experimental Economics*, Volume 11, pp. 122-133.

[29] Bartling, B., Engl, F. & Weber, R., 2014. Does willful ignorance deflect punishment? – An experimental study. *European Economic Review*, Volume 70, pp. 512-524.

[30] Baumeister, R., 1987. How the Self Became a Problem: A Psychological Review of Historical Research. *Journal of Personality and Social Psychology*, 52(1), pp. 163-176.

[31] Baumeister, R., 1998. The Self. In: D. Gilbert, S. Fiske & G. Lindzey, eds. *Handbook of Social Psychology.* New York: McGraw-Hill, pp. 680-740.

[32] Beckenkamp, M. et al., 2014. First Impressions are More Important than Early Intervention. Qualifying Broken Windows Theory in the Lab. *International Review of Law and Economics*, Volume 37, pp. 126-136.

[33] Becker, G., 1968. Crime and punishment: An economic approach. *Journal of Political Economy*, Volume 76, pp. 169-217.

[34] Becker, G., 1974. A theory of social interactions. *Journal of Political Economy*, 82(6), pp. 1063-1093.

[35] Becker, G. S., 1976. Altruism, egoism, and genetic fitness: economics and sociobiology. *Journal of Economic Literature*, 14(3), pp. 817-826.

[36] Becker, G. S., 1981. *A treatise on the family.* Boston: NBER Books.

[37] Bénabou, R. & Tirole, J., 2011. Identity, morals, and taboos: beliefs as assets. *Quarterly Journal of Economics*, 126(2), pp. 805-855.

[38] Bernheim, B. D., 1994. A Theory of Conformity. *Journal of Political Economy*, 102(5), pp. 841-877.

[39] Bernheim, B. D. & Exley, C. L., 2015. *Understanding Conformity: An Experimental Investigation,* s.l.: Harvard Business School Working Paper Series 16-070.

[40] Bicchieri, C., 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms.* New York: Cambridge University Press.

[41] Bicchieri, C., Lindemans, J. & Jiang, T., 2014. A structured approach to a diagnostic of collective practices. *Frontiers in Psychology*, Volume 5, pp. 1-13.

[42] Bicchieri, C. & Xiao, E., 2009. Do the Right Thing: But Only if Others do so. *Journal of Behavioral Decision Making*, 22(2), pp. 191-208.

[43] Bikhchandani, S., Hirshleifer, D. & Welch, I., 1998. Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *The Journal of Economic Perspectives*, 12(3), p. 151–170.

[44] Bikhchandani, S. & Sharma, S., 2000. *Herd Behavior in Financial Markets: A Review,* Washington D.C.: IMF Working Paper No. 00/48.

[45] Binmore, K. G., 1998. *Just Playing: Game Theory and the Social Contract.* Cambridge: MIT Press.

[46] Blume, L. E., Brock, W. A., Durlauf, S. N. & Ioannides, Y. M., 2010. Identification of Social Interactions. In: J. Benhabib, A. Bisin & M. O. Jackson, eds. *Social Economics.* s.l.:North Holland.

[47] Bogardus, E., 1928. *Immigration and Race Attitudes.* Boston: Heath.

[48] Bohnet, I. & Frey, B. S., 1999. Social Distance and other-regarding Behavior in Dictator Games: Comment. *The American Economic Review*, 89(1), pp. 335-339.

[49] Bolton, G. & Ockenfels, A., 2000. "ERC–A Theory of Equity, Reciprocity, and Competition. *American Economic Review*, 90(1), pp. 166-193.

[50] Bowles, S. & Gintis, H., 2004. The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical population biology*, 65(1), pp. 17-28.

[51] Boyd, R. & Richerson, P. J., 2001. Norms and bounded rationality. In: G. Gigerenzer & R. Selten, eds. *Bounded rationality: the adaptive toolbox.* s.l.:MIT Press, pp. 281-296.

[52] Brandts, J. & Solà, C., 2010. Personal relations and their effect on behavior in an organizational setting: an experimental study. *Journal of Economic Behavior and Organization*, Volume 73, pp. 246-253.

[53] Burke, M. A. & Sass, T. R., 2013. Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, 31(1), pp. 51-82.

[54] Bursztyn, L., Ederer, F., Ferman, B. & Yuchtman, N., 2014. Understanding Mechanisms Underlying Peer Effects: Evidence From a Field Experiment on Financial Decisions. *Econometrica*, 82(4), pp. 1273-1301.

[55] Cabral, L., 2005. *The Economics of Trust and Reputation: A Primer,* Mimeo: NYU.

[56] Cameron, L., Chaudhuri, A., Erkal, N. & Gangadharan, L., 2009. Propensities to Engage in and Punish Corrupt Behavior: Experimental Evidence from Australia, India, Indonesia and Singapore. *Journal of Public Economics*, Volume 93, pp. 843-851.

[57] Carrillo, J. & Mariotti, T., 2000. Strategic ignorance as a self-disciplining device. *Review of Economic Studies*, Volume 67, pp. 529-544.

[58] Case, A. & Katz, L., 1991. *The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths,* s.l.: NBER Working Paper No. 3705.

[59] Chari, V. & Kehoe, P., 2004. Financial Crises as Herds: Overturning the Critiques. *Journal of Economic Theory*, 119(1), pp. 128-150.

[60] Charness, G. & Dufwenberg, M., 2006. Promises and partnership. *Econometrica*, Volume 74, pp. 1579-1601.

[61] Charness, G. & Gneezy, U., 2008. What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization*, Volume 68, pp. 29-35.

[62] Charness, G. & Rabin, M., 2002. Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics*, Volume 117, pp. 817-869.

[63] Chaudhuri, A., 2011. Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. *Experimental Economics*, Volume 14, pp. 47-83.

[64] Chen, R. & Chen, Y., 2011. The potential of social identity for equilibrium selection. *American Economic Review*, 101(6), p. 2562–2589.

[65] Chen, Y. & Li, S., 2009. Group Identity and Social Preferences. *American Economic Review*, 99(1), pp. 431-457.

[66] Chetty, R., Hendren, N. & Katz, L. F., 2015. *The Effects of Exposure to Better Neighborhoods on Children: New Evidence From the Moving To Opportunity Experiment,* s.l.: NBER Working Paper 21156.

[67] Cho, I.-K. & Kreps, D., 1987. Signaling games and stable equilibria. *Quarterly Journal of Economics*, Volume 102, pp. 179-221.

[68] Cialdini, R. B. & DeNichols, M. E., 1989. Self-presentation by association. *Journal of Personality and Social Psychology*, Volume 57, pp. 626-631.

[69] Cialdini, R. B. & Goldstein, N. J., 2004. Social Influence: Compliance and Conformity. *Annual Review of Psychology*, Volume 55, pp. 591-621.

[70] Cialdini, R., Reno, R. & Kallgren, C., 1990. A Focus Theory Of Normative Conduct: Recycling The Concept Of Norms To Reduce Littering in Public Places. *Journal of Personality and Social Psychology*, pp. 1015-1026.

[71] Conlisk, J., 1980. Costly Optimizers versus Cheap Imitators. *Journal of Economic Behavior and Organization*, Volume 1, pp. 275-293.

[72] Conrads, J. & Irlenbusch, B., 2013. Strategic ignorance in ultimatum bargaining. *Journal of Economic Behavior & Organization*, Volume 92, pp. 104-115.

[73] Cox, J., Friedman, D. & Sadiraj, V., 2008. Revealed altruism. *Econometrica*, 76(1), pp. 31-69.

[74] Croson, R. & Shang, J., 2008. The Impact of Downward Social Information on Contribution Decisions. *Experimental Economics*, Volume 11, pp. 221-233.

[75] Dahl, G., Løken, K. & Mogstad, M., 2014. Peer effects in program participation. *American Economic Review*, 104(7), p. 2049–2074.

[76] Damm, A. P. & Dustmann, C., 2014. Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?. *American Economic Review*, 104(6), pp. 1806-1832.

[77] Dana, J., Weber, R. A. & Kuang, J. X., 2007. Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, Volume 33, pp. 67-80.

[78] Dawes, R., 1980. Social Dilemmas. *Annual Review of Psychology*, 31(2), pp. 169-193.

[79] DellaVigna, S., List, J. & Malmendier, U., 2012. Testing for altruism and social pressure in charitable giving. *Quarterly Journal of Economics*, 127(1), pp. 1-56.

[80] Dimant, E. & Schulte, T., forthcoming. The nature of corruption: An interdisciplinary perspective. *German Law Journal*.

[81] Eckel, C. C. & Grossman, P. J., 1996. Altruism in anonymous dictator games. *Games and Economic Behavior*, 16(2), pp. 181-191.

[82] Eckel, C. & Grossman, P., 2005. Managing diversity by creating team identity. *Journal of Economic Behavior and Organization*, 58(3), pp. 371-392.

[83] Eckel, C. & Petrie, R., 2011. Face Value. *American Economic Review*, 101(4), pp. 1497-1513.

[84] Ellingsen, T. & Johannesson, M., 2008. Pride and prejudice: the human side of incentive theory. *American Economic Review*, 98(3), pp. 990-1008.

[85] Ellison, G. & Fudenberg, D., 1993. Rules of Thumb for Social Learning. *Journal of Political Economy*, 101(4), p. 612–643.

[86] Elster, J., 1989. *The Cement of Society – A Study of Social Order.* Cambridge: Cambridge University Press.

[87] Evans, W., Oates, W. & Schwab, R., 1992. Measuring Peer Group Effects: A Study of Teenage Behavior. *The Journal of Political Economy*, 100(5), pp. 966-991.

[88] Falk, A. & Fischbacher, U., 2002. "Crime" in the Lab - Detecting Social Interaction. *European Economic Review*, Volume 46, pp. 859-869.

[89] Falk, A. & Ichino, A., 2006. Clean Evidence on Peer Pressure. *Journal of Labor Economics*, Volume 24, pp. 39-57.

[90] Fehr, E. & Fischbacher, U., 2005. The Economics of Strong Reciprocity. In: H. Gintis, S. Bowles, R. Boyd & E. Fehr, eds. *Moral Sentiments and Material Interests. The Foundations for Cooperation in Economic Life.* Cambridge: MIT Press, pp. 151-193.

[91] Fehr, E. & Gächter, S., 2000. Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, Volume 90, pp. 980-994.

[92] Fehr, E. & Schmidt, K., 1999. A Theory of Fairness, Competition and Cooperation. *The Quarterly Journal of Economics*, 114(3), pp. 817-868.

[93] Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), pp. 171-178.

[94] Frey, B., 1997. *Not Just For the Money – An Economic Theory of Personal Motivation.* Cheltenham: Edward Elgar Publishing Limited.

[95] Frey, B. & Meier, S., 2004. Social Comparison and Pro-Social Behavior: Testing Conditional Cooperation in a Field Experiment. *American Economic Review*, 94(5), p. 1717–1722.

[96] Gächter, S., Kölle, F. & Quercia, S., 2015. *Different frames or different games? Comparing common pool and public good social dilemmas,* University of Nottingham: Mimeo.

[97] Gächter, S., Nosenzo, D. & Sefton, M., 2012. The Impact of Social Comparisons on Reciprocity. *The Scandinavian Journal of Economics*, 114(4), pp. 1346-1367.

[98] Gächter, S., Nosenzo, D. & Sefton, M., 2013. Peer Effects in Pro-Social Behavior: Social Norms or Social Preferences. *Journal of the European Economic Association*, 11(3), pp. 548-573.

[99] Galinsky, A. D., Ku, G. & Wang, C. S., 2005. Perspective-taking: increasing social bonds and facilitating social coordination. *Group Processes and Intergroup Relations*, Volume 8, pp. 109-124.

[100] Georganas, S., Tonin, M. & Vlassopoulos, M., 2013. *Peer pressure and productivity: the role of observing and being observed,* s.l.: IZA DP No. 7523.

[101] Gibbs, J. L., Ellison, N. B. & Heino, R. D., 2006. Self-presentation in online personals: the role of anticipated future interaction, self-disclosure, and perceived success in internet dating. *Communication Research*, 33(2), pp. 152-177.

[102] Gino, F., Ayal, S. & Ariely, D., 2009. Contagion and Differentiation in Unethical Behavior. *Psychological Science*, 20(3), pp. 393-398.

[103] Gino, F. & Bazerman, M. H., 2009. When misconduct goes unnoticed: the acceptability of gradual erosion in others' unethical behavior. *Journal of Experimental Social Psychology*, Volume 45, p. 708–719.

[104] Gino, F. & Galinsky, A. D., 2012. Vicarious dishonesty: when psychological closeness creates distance from one's moral compass. *Organizational Behavior and Human Decision Processes*, Volume 119, pp. 15-26.

[105] Glaeser, E. L., Sacerdote, B. I. & Scheinkman, J. A., 2003. The Social Multiplier. *Journal of the European Economic Association*, 1(2-3), pp. 345-353.

[106] Glaeser, E. L. & Scheinkman, J. A., 2004. Measuring Social Interactions. In: S. N. Durlauf & H. P. Young, eds. *Social Dynamics (Economic Learning and Social Evolution Series).* s.l.:MIT University Press Group Ltd., pp. 83-131.

[107] Glaeser, E., Sacerdote, B. & Scheinkman, J., 1996. Crime and Social Interaction. *The Quarterly Journal of Economics*, 111(2), pp. 507-548.

[108] Gneezy, U., 2005. Deception: the role of consequences. *American Economic Review*, Volume 95, pp. 384-394.

[109] Goethe, J. W., 1929. *The sorrows of young Werther (W. Rose, translated).* London: Scholastic Press.

[110] Goldstein, N., Cialdini, R. & Griskevicius, V., 2008. A room with a viewpoint: using social norms to motivate environmental conservation in Hotels. *Journal of Consumer Research*, Volume 35, p. 472–482.

[111] Goldstein, N. J. & Cialdini, R. B., 2007. The spyglass self: a model of vicarious self-perception. *Journal of Personality and Social Psychology*, Volume 92, pp. 402-417.

[112] Gould, E. & Kaplan, T., 2011. Learning unethical practices from a co-worker: the peer effect of Jose Canseco. *Labour Economics*, 18(3), pp. 338-348.

[113] Granovetter, M., 1979. Threshold Models of Collective Behavior. *American Journal of Sociology*, Volume 83, pp. 1420-1443.

[114] Greiner, B., 2004. An online recruitment system for economic experiments. *Forschung und Wissenschaftliches Rechnen*, Volume 63, pp. 79-93.

[115] Grossman, Z., 2010. *Self-Signaling Versus Social-Signaling in Giving.,* University of Santa Barbara: WP.

[116] Grossman, Z., 2014. Strategic ignorance and the robustness of social preferences. *Management Science*, 60(11), pp. 2659-2665.

[117] Gunia, B. C., Sivanathan, N. & Galinsky, A. D., 2009. Vicarious entrapment: Your sunk costs, my escalation of commitment. *Journal of Experimental Social Psychology*, 45(6), pp. 1238-1244.

[118] Herrmann, B., Thöni, C. & Gächter, S., 2008. Antisocial punishment across societies. *Science*, Volume 319, pp. 1362-1367.

[119] Hitsch, G., Hortaçsu, A. & Ariely, D., 2010. Matching and sorting in online dating. *American Economic Review*, 100(1), pp. 130-163.

[120] Hoffman, E., McCabe, K. & Smith, V. L., 1996. Social distance and other regarding behavior in dictator games. *American Economic Review*, 86(3), pp. 653-660.

[121] Hogg, M. & Tindale, S., 2002. *Blackwell Handbook of Social Psychology: Group Processes.* s.l.:Wiley-Blackwell.

[122] Houser, D., Vetter, S. & Winter, J., 2012. Fairness and Cheating. *European Economic Review*, 56(8), pp. 1645-1655.

[123] Hyman, H., 1942. The Psychology of Status. *Archives of Psychology*, Issue 269.

[124] Ichino, A. & Maggi, G., 2000. Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm. *The Quarterly Journal of Economics*, 115(3), pp. 1057-1090.

[125] Innes, R. & Mitra, A., 2013. Is dishonesty contagious?. *Economic Inquiry*, 51(1), pp. 722-734.

[126] Irlenbusch, B. & Villeval, M., 2015. Behavioral ethics: how psychology influenced economics and how economics might inform psychology?. *Current Opinion in Psychology*, Volume 6, pp. 87-92.

[127] Jankowski, M., 1991. *Islands in the Street - Gangs and American Urban Society.* Berkeley: University of California Press.

[128] Jones, S., 1984. *The Economics of Conformism.* Oxford: Basil Blackwell.

[129] Kahneman, D. & Tversky, A., 2009. *Choices, Values and Frames.* 10 ed. Cambridge: Cambridge University Press.

[130] Kandel, E. & Lazear, E., 1992. Peer pressure and partnerships. *Journal of Political Economy*, Volume 100, pp. 801-817.

[131] Katz, L. F., Kling, J. R. & Liebman, J. B., 2001. Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment. *The Quarterly Journal of Economics*, 116(2), pp. 607-654.

[132] Kelling, G. & Wilson, J., 1982. Broken Windows. *The Atlantic,* pp. 1-11.

[133] Kelman, H., 1958. Compliance, Identification, and Internalization: Three Processes of Attitude Change. *Journal of Conflict Resolution*, 2(1), pp. 51-60.

[134] Keynes, J., 1936. *The General Theory of Employment, Interest and Money.* London: Macmillan.

[135] Khadjavi, M., 2015. On the interaction of deterrence and emotions. *The Journal of Law, Economics & Organization*, 31(2), pp. 287-319.

[136] Kling, J. R., Liebman, J. B. & Katz, L. F., 2007. Experimental Analysis of Neighborhood Effects. *Econometrica*, 75(1), pp. 83-119.

[137] Kling, J. R., Ludwig, J. & Katz, L. F., 2005. Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment. *The Quarterly Journal of Economics*, 120(1), pp. 87-130.

[138] Krugman, P. R., 1991. Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3), pp. 483-499.

[139] Lahno, A. & Serra-Garcia, M., 2015. Peer Effects in Risk Taking. *Journal of Risk and Uncertainty*, Volume 50, pp. 73-95.

[140] Laland, K. N., 2002. Imitation, Social Learning, and Preparedness as Mechanisms of Bounded Rationality. In: G. Gigerenzer & R. Selten, eds. *Bounded rationality: the adaptive toolbox.* s.l.:MIT Press, pp. 233-247.

[141] Laundry, C. et al., 2006. Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment. *Quarterly Journal of Economics*, 121(2), p. 747–782.

[142] Lefebvre, M., Pestieau, P., Riedl, A. & Villeval, M., 2015. Tax evasion and social information: an experiment in Belgium, France and the Netherlands. *International Tax and Public Finance*, Volume 22, pp. 401-425.

[143] Leider, S., Möbius, M. M., Rosenblat, T. & Do, Q.-A., 2009. Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics*, 124(4), pp. 1815-1851.

[144] Levitt, S. D., 1999. The exaggerated role of changing age structure in explaining aggregate crime changes. *Criminology*, 37(3), pp. 581-597.

[145] Levitt, S. & List, J., 2009. Field experiments in economics: The past, the present, and the future. *European Economic Review*, 53(1), pp. 1-18.

[146] List, J., 2007. On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3), pp. 482-493.

[147] Lorenz, E. N., 1963. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20(2), pp. 130-141.

[148] Ludwig, J. & Kling, J. R., 2007. Is Crime Contagious?. *Journal of Law and Economics*, 50(3), pp. 491-518.

[149] Luhtanen, R. & Crocker, J., 1992. A collective self-esteem scale: self-evaluation of one's social identity. *Personality and social psychology bulletin*, 18(3), pp. 302-318.

[150] Manski, C., 1993. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, Volume 60, pp. 531-542.

[151] Manski, C., 2000. Economic Analysis of Social Interactions. *The Journal of Economic Perspectives*, 14(3), pp. 115-136.

[152] Mas, A. & Moretti, E., 2009. Peers at Work. *American Economic Review*, 99(1), pp. 112-145.

[153] Mazar, N., Amir, O. & Ariely, D., 2008. The dishonesty of honest people: a theory of self-concept maintenance. *Journal of Marketing Research*, 45(6), pp. 633-644.

[154] McAdams, R. H., 1997. The origin, development, and regulation of norms. *Michigan Law Review*, Volume 96, pp. 338-443.

[155] McLeod, J. M., Price, K. O. & Harburg, E., 1966. Socialization, liking and yielding of opinions in imbalanced situations. *Sociometry*, Volume 29, pp. 197-212.

[156] Merton, R., 1957. *Social Theory and Social Structure.* Glencoe: The Free Press.

[157] Milgram, S., 1963. Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, 67(4), pp. 371-378.

[158] Milgram, S., 1974. *Obedience to Authority: An Experimental View.* New York: Harper Row. Publishing.

[159] Mittone, L. & Ploner, M., 2011. Peer Pressure, Social Spillovers, and Reciprocity: An Experimental Analysis. *Experimental Economics*, Volume 14, pp. 203-222.

[160] Mulligan, C. B., 1998. Pecuniary incentives to work in the United States during World War II. *Journal of Political Economy*, 106(5), pp. 1033-1077.

[161] Murphy, K. M., Shleifer, A. & Vishny, R., 1993. Why is Rent Seeking so Costly to Growth?. *American Economic Review*, 83(2), pp. 409-414.

[162] National Institute of Justice, 2014. *Recidivism.* [Online]
Available at: http://www.nij.gov/topics/corrections/recidivism/pages/welcome.aspx
[Accessed 21 08 2015].

[163] Neumann, S., Paul, S. & Doering, P., 2013. *Remuneration and Risk-Taking - Empirical Evidence for Social Trading,* Bochum: MIMEO.

[164] Pelham, B., Carvallo, M. & Jones, J. T., 2005. Implicit egoism. *Current Directions in Psychological Science*, 14(2), pp. 106-110.

[165] Pennsylvania State University's Justice Center for Research, 2014. *Desistance from crime.* [Online]
Available at: www.justicecenter.psu.edu
[Accessed 21 08 2015].

[166] Phillips, D. P., 1974. The influence of suggestion on suicide: substantive and theoretical implications of the Werther effect. *American Sociological Review*, Volume 39, pp. 340-354.

[167] Pingle, M. & Day, R. H., 1996. Modes of economizing behavior: Experimental evidence. *Journal of Economic Behavior and Organization*, Volume 29, pp. 191-209.

[168] Ploner, M., 2013. Peer Effects at Campus Cafeterias - An Empirical Investigation Into Imitation and Social Proximity. *Journal of Evolutionary Economics*, Volume 23, pp. 61-76.

[169] Postlewaite, A., 2011. Social norms and preferences. In: J. Benhabib, A. Bisin & M. Jackson, eds. *Handbook of Social Economics.* s.l.:North Holland, pp. 31-67.

[170] Putnam, R., 2000. *Bowling Alone: The Collapse and Revival of American Community.* New York: Simon & Schuster Publishing.

[171] Rabin, M., 1993. Incorporating Fairness Into Game Theory and Economics. *American Economic Review*, 83(5), pp. 1281-1302.

[172] Rankin, A. M. & Philip, P. J., 1963. An epidemic of laughing in the Bukoba district of Tanganyika. *Central African Journal of Medicine*, Volume 9, p. 167–170.

[173] Rasmussen, E., 1996. Stigma and Self-Fulfilling Expectations of Criminality. *Journal of Law and Economics*, 39(2), pp. 519-543.

[174] Reiss, A. J., 1980. Understanding Changes in Crime Rates. In: S. Feinberg & A. J. Reiss, eds. *Indicators of Crime and Criminal Justice: Quantitative Studies.* Washington: Bureau of Justice Statistics.

[175] Reiss, A. J., 1988. Co-offending and Criminal Careers. In: M. Tonry & N. Morris, eds. *Crime and Justice - A Review of Research.* Chicago: University of Chicago Press.

[176] Richerson, P. J. & Boyd, R., 1998. The evolution of human ultra-sociality. In: I. Eibl-Eibisfeldt & E. Salter, eds. *Indoctrinability, Ideology, and Warfare: Evolutionary Perspectives.* New York: Berghahn, pp. 71-96.

[177] Roth, A., 1995. Bargaining experiments. In: J. Kagel & A. Roth, eds. *The Handbook of Experimental Economics.* Princeton: Princeton University Press, p. 253–348.

[178] Roth, A. E. & Erev, I., 1995. Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economics Behavior*, Volume 8, pp. 164-212.

[179] Sacerdote, B., 2001. Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics Economics*, 116(2), pp. 681-704.

[180] Salganik, M. J., Dodds, P. S. & Watts, D. J., 2006. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, Issue 311, pp. 854-856.

[181] Scharfstein, D. & Stein, J., 1990. Herd Behavior and Investment. *American Economic Review*, 80(3), pp. 465-479.

[182] Schelling, T., 1968. The life you save may be your own. In: S. Chase, ed. *Problems in public expenditure analysis.* Washington D.C.: Brookings Institution, pp. 127-162.

[183] Schlag, K., 1998. Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78(1), pp. 130-156.

[184] Selten, R., 1978. The chain store paradox. *Theory and Decision*, 9(2), pp. 127-159.

[185] Shang, J. & Croson, R., 2008. Field Experiments in Charitable Contribution: The Impact of Social Influence on the Voluntary Provision of Public Goods. *The Economic Journal*, 119(540), pp. 1422-1439.

[186] Shelling, T., 1973. Hockey Helmets, Concealed Weapons, and Daylight Saving. *Journal of Conflict Resolution*, 17(3), pp. 381-428.

[187] Sherif, M., 1936. *The Psychology of Social Norms.* New York: Harper Collins Publishing.

[188] Sherif, M., 1937. An Experimental Approach to the Study of Attitudes. *Sociometry*, 1(1), pp. 90-98.

[189] Shiller, R., 1984. Stock Prices and Social Dynamics. *Brookings Papers on Economic Activity*, Volume 2, pp. 457-498.

[190] Sinclair, P. J. N., 1990. The Economics of Imitation. *Scottish J. Pol. Econ.*, Volume 37, pp. 113-144.

[191] Sliwka, D., 2007. Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes. *American Economic Review*, 97(3), pp. 999-1012.

[192] Sutherland, E., 1939. *Principles of Criminology.* Philadelphia: Lipincott.

[193] Tajfel, H., 1982. *Social identity and intergroup relations.* Cambridge, England: Cambridge University Press.

[194] Tajfel, H. & Turner, J., 1979. An Integrative Theory of Intergroup Conflict. In: W. Austin & S. Worchel, eds. *The Social Psychology of Intergroup Relations.* Pacific Grove, CA: Brooks/Cole, pp. 33-47.

[195] Tajfel, H. & Turner, J., 1986. The social identity theory of intergroup behavior. In: S. Worchel & W. Austin, eds. *Psychology of Intergroup Relations.* Chicago: Nelson-Hall, p. 7–24.

[196] Tarde, G., 1903. *The Laws of Imitation.* New York: H. Holt and Company.

[197] Taylor, S. E., 1989. *Positive illusions: Creative self-deception and the healthy mind.* New York: Basic Books.

[198] Tenbrunsel, A. E. & Messick, D. M., 2004. Ethical fading: the role of self-deception in unethical behavior. *Social Justice Research*, 17(2), pp. 223-236.

[199] The New York Times, 2011. *Tax Evasion Costs Governments $3.1 Trillion Annually, Report Says.* [Online] Available at: http://www.nytimes.com/2011/11/26/business/global/26iht-tax26.html?_r=0 [Accessed 20 08 2015].

[200] The World Bank, 2013. *Six Questions on the Cost of Corruption with World Bank Institute Global Governance Director Daniel Kaufmann.* [Online] Available at: http://go.worldbank.org/KQH743GKF1 [Accessed 19 07 2013].

[201] Thöni, C. & Gächter, S., 2015. Peer Effects and Social Preferences in Voluntary Cooperation. *Journal of Economic Psychology*, Volume 48, pp. 72-88.

[202] Tirole, J., 1996. A Theory of Collective Reputations (with Applications to the Persistence of Corruption andto Firm Quality). *The Review of Economic Studies*, 63(1), pp. 1-22.

[203] United Nations Office on Drugs and Crime, 2011. *UNODC World Drug Report.* [Online] Available at: https://www.unodc.org/documents/data-and-analysis/WDR2011/World_Drug_Report_2011_ebook.pdf [Accessed 20 08 2015].

[204] Veldhuizen, R., Oosterbeek, H. & Sonnemans, J., 2014. *Peers at work: from the field to the lab,* s.l.: Tinbergen Institute Discussion Paper TI 2014-051/I.

[205] Wasserman, S. & Faust, K., 1994. *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press.

[206] Weber, M., 1978. *Economy and Society.* Berkeley: University of California Press.

[207] Wells, L., 1978. Theories of Deviance and the Self-Concept. *Social Psycology*, 41(3), pp. 189-204.

[208] Welsh, D. T., Ordóñez, L. D., Snyder, D. G. & Christian, M. S., 2015. The slippery slope: how small ethical transgressions pave the way for larger future transgressions. *Journal of Applied Psychology*, 100(1), pp. 114-127.

[209] Wheeler, L., 1966. Toward a Theory of Behavioral Contagion. *Psychological Review*, 73(2), pp. 179-192.

[210] Young, H. P., 2015. The evolution of social norms. *Annual Review of Economics*, Volume 7, pp. 359-387.

[211] Zafar, B., 2011. An Experimental Investigation of Why Individuals Conform. *European Economic Review*, Volume 55, pp. 774-798.

[212] Zimbardo, P. G., 1971. *The power and pathology of imprisonment,* s.l.: Congressional Record (Serial No. 15, Oct. 25, 1971) U.S. Government Printing Office.

[213] Zizzo, D. J., 2010. Experimenter demand effects in economic experiments. *Experimental Economics*, Volume 13, pp. 75-98.

10. Appendix

A: Data Overview, Robustness Checks, and Additional Results

	Treatments			Total
	Baseline (no Proximity)	Treatment 1 (High Proximity)	Treatment 2 (Low Proximity)	
Observed Ethical Behavior	44	27	34	105
Observed Unethical Behavior	39	42	41	122
Total	83	69	75	227

A1: Summary table of the observations by treatment variation and observed unethicality

Our theoretical foundation (chapter 5.2.2) assumes behavioral adaptation to be asymmetrically biased towards unethical behavior, holding social identification constant. Thus, for *adaptation gap (%)*, the results suggest that the observation of ethical (unethical) behavior leads to a less (more) pronounced closing of the adaptation gap. More precisely, the results indicate that observing unethical behavior leads to a revision behavior that reduces the gap between own initial and observed peer's behavior more strongly (from 100% down to 69.4%) as compared to observing ethical behavior (from 100% down to 78.5%). The differences are marginally significant at the 10% level ($p < 0.074$, $z = 1.787$). This implies that, if anything, individuals who observe unethical behavior go a long way and close the observed gap by some 66% through revising their initial decision, while individuals who observe ethical behavior reduce the gap only slightly by 14%. This result is not only in line with the theoretical predictions of our asymmetric model, but also suggests that one's inherent conquest to do good is less pronounced than to do bad, which yields additional support not only for hypothesis $\mathbf{H_1}$ but also for our asymmetric model specification.

A2: Adaptation gap (%) by observed (un)ethicality. The gap is calculated as the difference between the individual's revised decision and the behavior observed from the peer relative to the difference between one's initial behavior and the behavior observed from the peer. That is, the adaptation gap indicates how different the behavior between two participants are after the active participant was given the chance to revise his/her initial behavior. In this sense, the narrower the gap, the more similar the active and the respective passive participants are in terms of (un)ethical behavior. Participants who did not revise their initial behavior are treated as adaptation gap = 100%. The analysis is broken down into the (un)ethicality of observed behavior by the active participants

Overall, the results relating to hypothesis H_1 (see Figure 3 and A2) indicate that not only is the contagion of ethical behavior only half as likely as contagion of unethical behavior but it is also less pronounced. Following our results, ethical behavior is imitated only to a limited extent as compared to the imitation of unethical behavior.

We ran several OLS specifications to control for treatment effects, initial behavioral heterogeneity, observed behavior, and gender (including several interaction effects). Although not explicitly shown in the regression table, adding controls for, among others, risk, self-control, and greed show up as insignificant and did not alter the robustness of the results presented here. Overall, the presented results are robust to the inclusion of controls and across different specifications and estimation methods. Estimations are available upon request. In addition, we ran robustness checks using Tobit estimations in order to account for potentially censored behavior caused by ceiling effects in revision behavior.

The estimations yield a number of interesting results. For example, the high proximity condition, in which social identity was highest, shows up significantly negative across different specifications. The negative coefficient suggests that comparing to a situation in which social proximity remained unknown to the observer (Baseline), being exposed to a high proximity signal induced a downward revision of one's initial behavior. More precisely, participants took away 29%-36% more money from the charity in the high social identification condition independent of the (un)ethicality of the observed behavior. This effect is not statistically significantly pronounced for the low proximity condition. These numbers are in line with the findings presented previously, indicating that peer effects are particularly present in a high social identification environment, which over-proportionally crowds-out ethical behavior and leads to more unethical behavior. The adaptation gap observed by the individual, which is the difference between one's own and the peer's initial behavior in monetary terms, leads to a similar reasoning: individuals over-proportionally react to a larger gap by negatively revising one's initial amount, indicating that an increase in observed adaptation gap leads to more unethical behavior overall.

In addition, our results also suggest that one's initial behavior strongly predicts the direction of behavioral contagion. This speaks to the idea that the individual's predisposition to behave (un)ethically is essential for the direction of behavioral spillovers caused by peer effects. More precisely, behaving (un)ethically when peer effects are absent renders it likely to behave even more (un)ethically when being exposed to peer effects. Compared to the base level of no change in the initially fair 50:50 split between oneself and the charity, the revision of one's initial (un)ethical behavior is similar in either direction in terms of magnitude. Relative to one's initial behavior, the numbers indicate that an individual is likely to donate (take) between 9%-19% (19-23%) more money when his/her initial decision was to donate (take) money after

observing peer's behavior, thus again supporting the asymmetry of behavioral contagion. Consequently, beyond social identification one's initial behavior is highly predictive of how an individual reacts to peer effects. The results also suggest that initial behavior represents a trait that is consistent and robust to the exposure to peer effects: those who decided to behave (un)ethically in the first place are likely to remain (un)ethical, only to a more pronounced extent. We do not observe behavioral heterogeneity across gender. The table A3 below illustrates our results. In addition to analyzing the magnitude of behavioral change, we find the same robust results when looking at the drivers of behavioral contagion in general. Results using logit estimations are reported in Appendix A4. What is more, we find similarly robust results when further subdividing *No Contagion* into those who were invariant to the observed behavior and thus did not react at all and those who reacted but changed their behavior into the opposite direction of what they have observed (anti-contagion). We apply a multinomial logit regression and report results in Appendix A5.

Dependent Variable: Amount Revised (%)	OLS Specifications				Tobit Specifications			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatments *(Base Level: Unknown Proximity)*								
High Proximity	-29.3860*		-28.7364*	-30.7496**	-34.9762***		-34.2641***	-36.0689***
	(16.3469)		(16.5249)	(14.6992)	(9.7107)		(9.6609)	(9.4863)
Low Proximity	-6.1454		-5.1774	-6.4478	-6.8891		-5.9119	-6.9097
	(4.4392)		(4.5562)	(5.1175)	(6.2096)		(6.2084)	(6.4023)
Observed Unethicality		37.0772	34.0409	34.2158		38.1142	34.5891	33.7526
		(24.2322)	(25.8306)	(39.6806)		(24.6067)	(24.0454)	(26.7849)
Initial Behavior *(Base Level: No Change of Fair Split)*								
Give	9.2353**	15.1657**	15.5747**	18.6116**	10.0438	15.9512*	16.4712*	19.3937**
	(4.5455)	(7.0586)	(6.9108)	(9.3087)	(7.3871)	(8.8428)	(8.5974)	(8.7559)
Take Away	-9.6093	-21.4052***	-19.6451***	-18.7158*	-11.1630	-23.3327**	-21.3697**	-20.5225**
	(6.1702)	(7.9535)	(7.4600)	(10.1752)	(7.0013)	(10.1893)	(9.9465)	(9.9621)
Observed Behavioral Gap	-0.0436***	-0.1022**	-0.0978*	-0.1085	-0.0471***	-0.1070***	-0.1022**	-0.1102**
	(0.0121)	(0.0480)	(0.0510)	(0.0756)	(0.0096)	(0.0405)	(0.0395)	(0.0428)
Number of Interests				-1.9073				-1.8053
				(4.1709)				(3.9709)
Gender				-7.5626				-7.8040
				(15.0804)				(15.7833)
Interaction 1 *(Gender x Observed Unethicality)*				-5.1273				-4.9789
				(26.6440)				(25.8968)
Interaction 2 *(Gender x Observed Behavioral Gap)*				0.0345				0.0318
				(0.0548)				(0.0410)
Constant	0.7803	-22.9721*	-16.3068	26.1370	1.0529	-24.0145*	-16.2959	-13.1646
	(3.3260)	(11.8731)	(14.1127)	(24.8911)	(5.7302)	(12.9784)	(13.3357)	(14.5095)
Dummies Questionnaire	No	No	No	Yes^{NS}	No	No	No	Yes^{NS}
Observations	185	185	185	185	185	185	185	185
Adjusted R^2	0.138	0.105	0.144	0.137				

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

A3: Amount Revised (%) as a function of social identification and observed unethicality. Model 1 and 5 (Model 2 and 6) tests for the effect of the treatments (of observed unethicality) while controlling for initial behavior and the observed adaptation gap. Model 3 (Model 7) tests for both treatment effects and observed unethicality simultaneously. Model 4 (Model 8) adds controls for gender and some interaction terms as well as the number of interest. In order to rule out any endogenous concerns and stress the effectiveness of exogenous variation of social proximity we also add dummies for the dating website questions. None of the dummies turn out significant in neither model specification, thus emphasizing the robustness of our results.

Dependent Variable: Adaptation Gap (%)	**Logit Specifications**			
	(1)	(2)	(3)	(4)
Treatments				
(Base Level: Unknown Proximity)				
High Proximity	35.4372***		35.0685***	54.4839***
	(25.4433)		(25.2349)	(50.1293)
Low Proximity	7.7127***		7.5854***	16.4058***
	(4.4677)		(4.4276)	(12.1507)
Observed Unethicality		0.2577	0.6840	0.5725
		(0.3718)	(1.1760)	(1.5174)
Initial Behavior				
(Base Level: No Change of Fair Split)				
Give	1.3153	1.0835	1.2070	1.0064
	(0.7233)	(0.6430)	(0.8143)	(0.8653)
Take Away	2.9752**	4.4048**	3.2950*	3.8362
	(1.5292)	(2.6915)	(2.2759)	(3.7619)
Observed Behavioral Gap	1.0015**	1.0035	1.0021	1.0030
	(0.0007)	(0.0024)	(0.0028)	(0.0043)
Number of Interests				0.9046
				(0.3099)
Gender				1.6828
				(2.7092)
Interaction 1				2.6033
(Gender x Observed Unethicality)				(6.5987)
Interaction 2				1.0006
(Gender x Observed Behavioral Gap)				(0.0043)
Dummies Questionnaire	No	No	No	Yes
Observations	184	184	184	184
AIC	157.0341	188.9185	158.9854	191.0477
BIC	176.0228	204.7425	181.1389	304.9800

Standard errors in parentheses. Odds ratios reported.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

A4: Examination of the adaptation gap (%) using Logit estimations. Participants who did not revise their initial decision are treated as adaptation gap = 100%.

	Dependent Variable: Conformity	**Logit Specifications**		
		(1)	(2)	(3)
Anti-Contagion (Repulsion)	**Treatments** *(Base Level: Unknown Proximity)*			
	High Proximity	1.7488* (1.0053)		1.7638* (1.0125)
	Low Proximity	0.2982 (0.7963)		0.3264 (0.8472)
	Observed Unethicality		0.8589 (3.1571)	0.8794 (3.3638)
	Initial Behavior *(Base Level: No Change of Fair Split)*			
	Give	1.5038* (0.8359)	1.7036 (1.0740)	1.6616 (1.0839)
	Take Away	-0.4926 (1.3722)	-0.8689 (2.2678)	-0.7659 (2.2235)
	Observed Behavioral Gap	-0.0000 (0.0013)	-0.0017 (0.0059)	-0.0014 (0.0059)
No Contagion (Invariance)	*Base Outcome*			
Contagion	**Treatments** *(Base Level: Unknown Proximity)*			
	High Proximity	3.6437*** (0.7034)		3.6327*** (0.7058)
	Low Proximity	2.0636*** (0.5895)		2.0453*** (0.5877)
	Observed Unethicality		-1.2356 (1.2851)	-0.5957 (1.6213)
	Initial Behavior *(Base Level: No Change of Fair Split)*			
	Give	0.3060 (0.5757)	0.1027 (0.6172)	0.1763 (0.7039)
	Take Away	1.1390** (0.5191)	1.4101** (0.5795)	1.2959** (0.6325)
	Observed Behavioral Gap	0.0016** (0.0007)	0.0032 (0.0022)	0.0025 (0.0027)
	Observations	184	184	184
	AIC	232.4102	264.1861	236.1646
	BIC	270.9894	296.3355	281.1737

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A5: Examination of the drivers of behavioral contagion using a multinomial logistic regression. Behavior of participants who revised their initial decision but into the opposite direction of what they have observed from their peer is treated as anti-contagion (repulsion).

Following the regression results, our results also speak to the idea that behavioral contagion seems to facilitate the magnitude of (un)ethical behavior rather than changing individual behavior to the better or worse, respectively. That is, those who behaved (un)ethically in the first place become even more (un)ethical after being exposed to their peers. This is true for both revision of one's initial amount and the reduction of the adaptation gap to one's peer. In more detail and in line with our previous results, this relation is more pronounced the more salient social identity to one's peers is and in particular in the unethical domain. This is additional support for our hypothesis H_1 that unethical behavior is more contagious than ethical behavior. We find similar results for those who decided to keep the fair equal split in the beginning, while the behavioral change of those who donated initially seem not to be susceptible to changes in social identity. The figure below illustrates results.

A6: The upper (lower) graph depicts the amount revised (%) (adaptation gap (%)) conditional on initial behavior broken down by social proximity. This figure depicts the percentage of initial behavior revised as a function of the participant's initial behavior and broken down into different social identity categories. Essentially, the figure illustrates the magnitude and direction of behavioral contagion after observing the passive peer conditional on one's initial decision prior to having observed a peer.

B: Social Identity Statements

1. I am a reliable person.
2. I am interested in politics and/or economics.
3. Money is important to me.
4. I am an honest and sincere person.
5. I am a cinephile.
6. I am interested in sports.
7. I am a religious person and faith is important to me.
8. I am fond of animals.
9. I am interested in art and/or cultures.
10. I am an active and adventurous person.
11. I am interested in cars and/or technology.
12. I am fond of children and family-oriented.
13. I am interested in foreign languages and/or countries.
14. I am a warmhearted and helpful person.
15. I am a tolerant person.
16. I like to gossip.
17. I am a faithful person.
18. I play an instrument.
19. I like to go out and dance.
20. I am a goal-oriented person.
21. I spend a lot of time in front of the TV.
22. I am a sociable person and like to be among people.
23. I like to play videogames.
24. I am a humorous and entertaining person.
25. I am a strong-willed person.

Average amount of chosen statements (across all treatments): 15.8 (63%)

C: Experimental Instructions

General Information on the Experiment

- First of all, we would like to thank you very much for participating in this experiment. Please read the instructions carefully. The experiment will last for about 45-60 minutes.

- During the entire experiment, no communication is allowed. If there is something you do not understand or if you have any questions, now or at some point during the experiment, please raise your hand and remain seated. One of our colleagues will come to you and answer your question.

- During the experiment, you have the possibility to earn money. The amount you will receive at the end of the session depends on how many "Taler" you earn during the experiment.

- At the end of the experiment, the amount of "Taler" that you have earned will be converted into real money at an exchange rate of 20 Taler = 1 Euro.

- All decisions you make during this experiment will remain anonymous. None of the participants gets to know the identity of other participants in the experiment and decisions cannot be linked to a specific participant. Moreover, you will be paid anonymously at the end of the experiment.

Order of Events:

- The experiment consists of a list of statements that you will receive at the beginning and further decisions. Explanations and information related to these decisions will be given as the experiment progresses. You will make these decisions **once**.

- Both **you** as well as a **charitable organization of your choice** (i.e. an officially registered charity organization) will be provisionally assigned a monetary amount of **300 Taler** each.

- During the experiment you will have to decided on whether you want to…
  - … *take* a part or all of the money from the charitable organization.
  - … *leave* the division of the sum of money as it is.
  - … *give* a part or all of your money to the charitable organization.

- In case you decide to **take money** from the charitable organization, the respective amount of money will be transferred to your individual cash account and exactly the same amount will be deducted from the cash account of the charitable organization.

- Should you decide to **give** money to the charitable organization of your choice, the respective amount of money will be deducted from your individual cash account and given to the charity. The experimenter will double all ECUs remaining in the charity's account at the end of the experiment.

- Your decision remains anonymous and neither the other participants of the experiment nor the experimenters have the possibility to assign your choices to your identity.

- **At the end** of the experiment, **one** participant will be chosen at random and his or her choice will be implemented and count towards the charity (i.e. that choice will be relevant for the payment). In particular, we will double the respective amount and donate it to the charity after the experiment ends. The receipt of this donation will be published on the homepage of the BaER-Lab (www.baer-lab.org) in a timely manner. All

other participants will receive **150 Taler** (including the show-up fee) at the end of the experiment.

- The total payoff of the participants:
  - **In case you are the randomly chosen participant**

    300 Taler +/- the amount of money that has been given to/taken from the cash account of the charitable organization
  - **In case you are not the randomly chosen participant**

    150 Taler

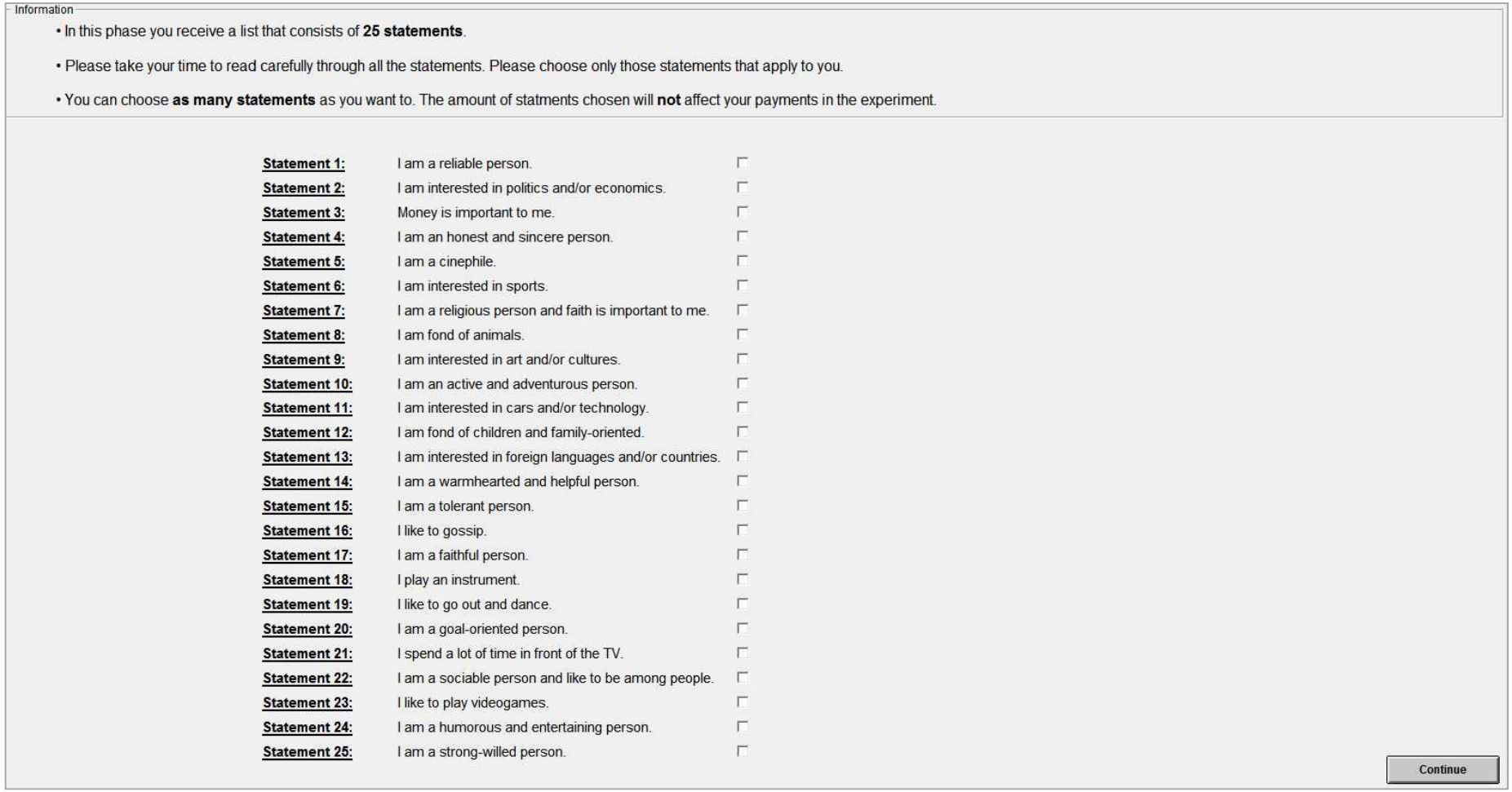
- The total payoff of the charitable organization:
  - (Amount of money in the cash account of the charitable organization of the randomly chosen participant) × 2

- At the end of the experiment, the relevant information on the payment will be made visible to each participant on his or her screen.

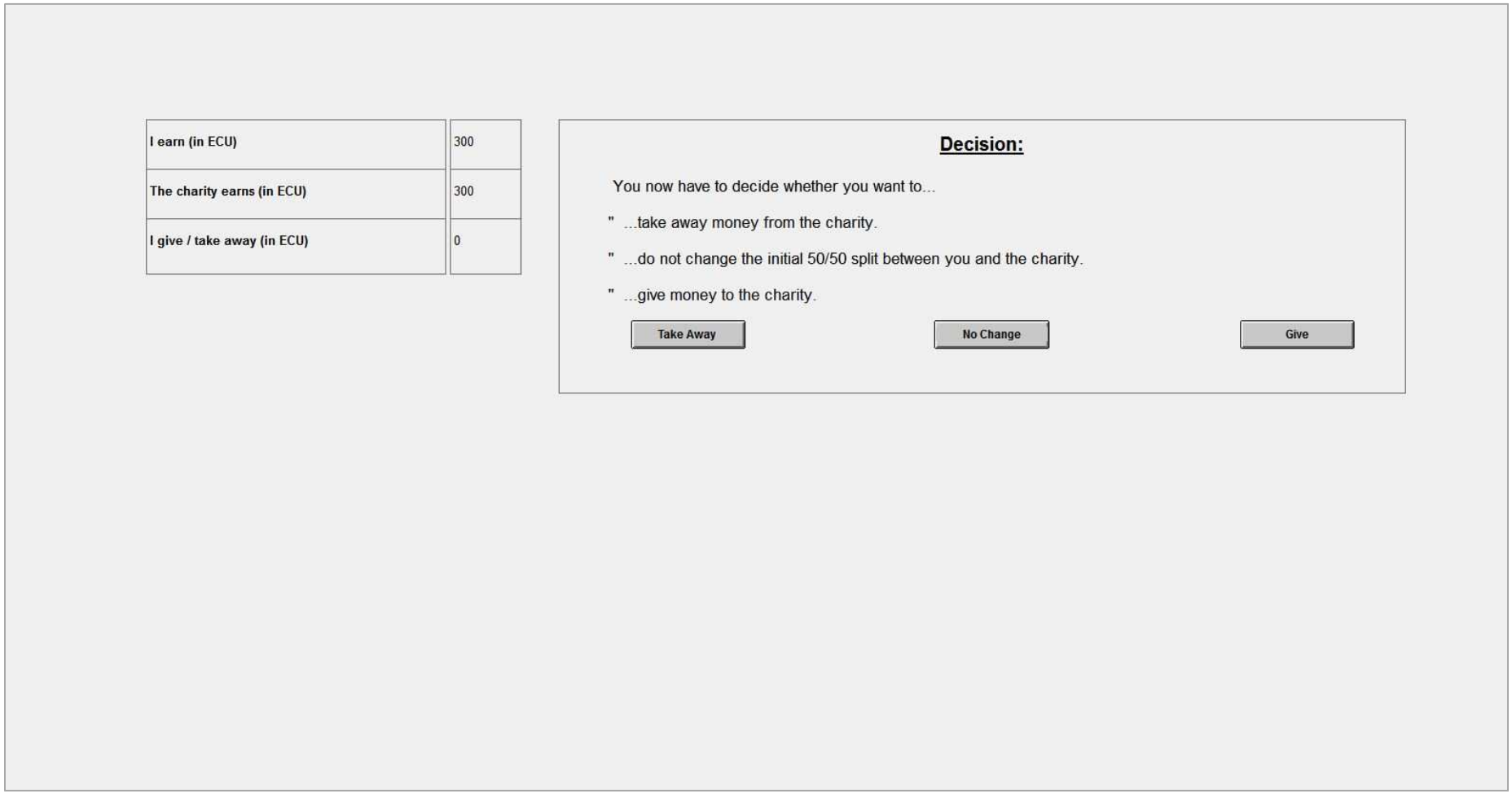
- After the actual experiment concludes, we will ask you to fill out a questionnaire. Please fill out the questionnaire carefully and truthfully.

D: Screenshots of Decision Screens

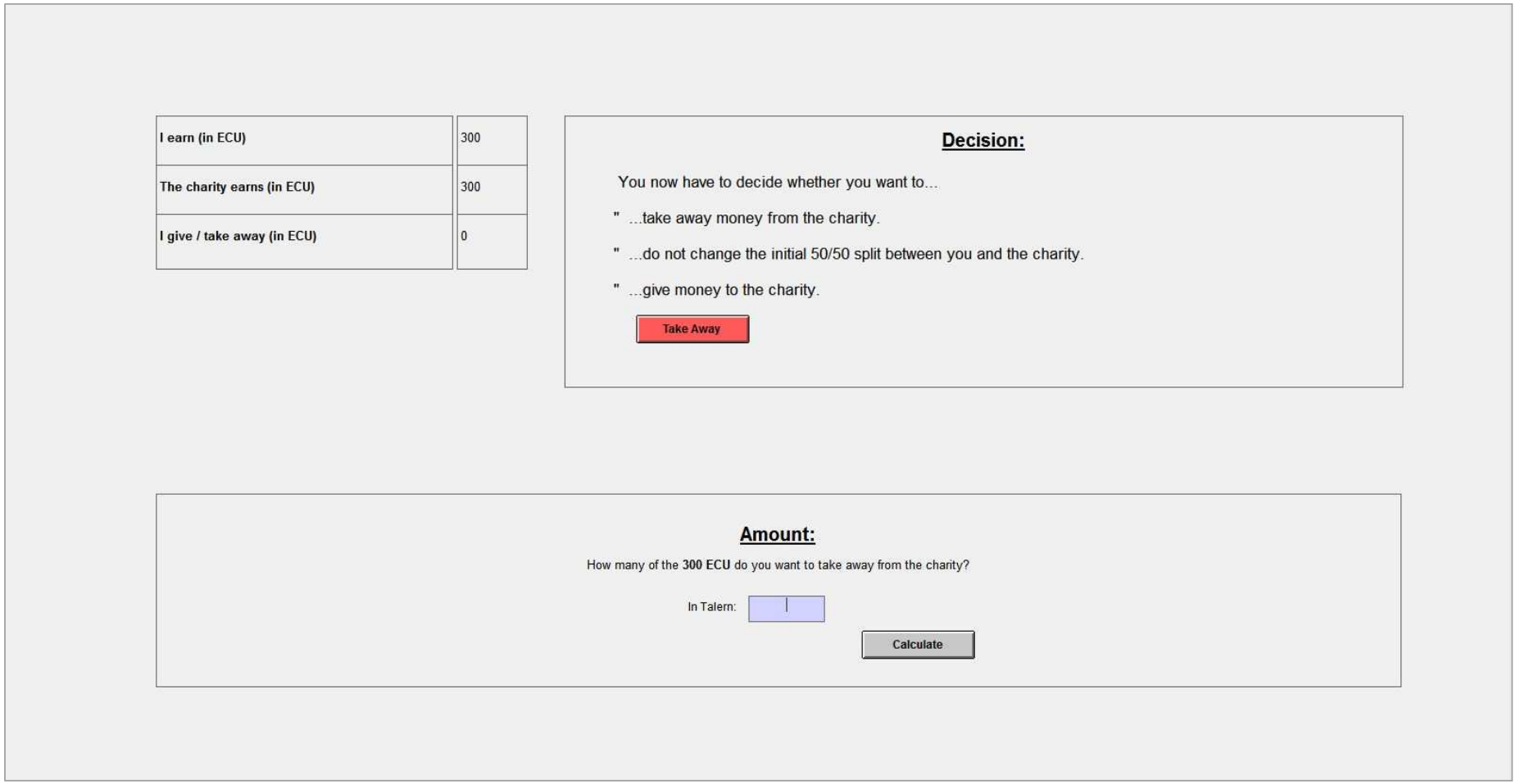
1. List of statements: generates the proximity measure in all treatments



2. First decision: behaving (un)ethically

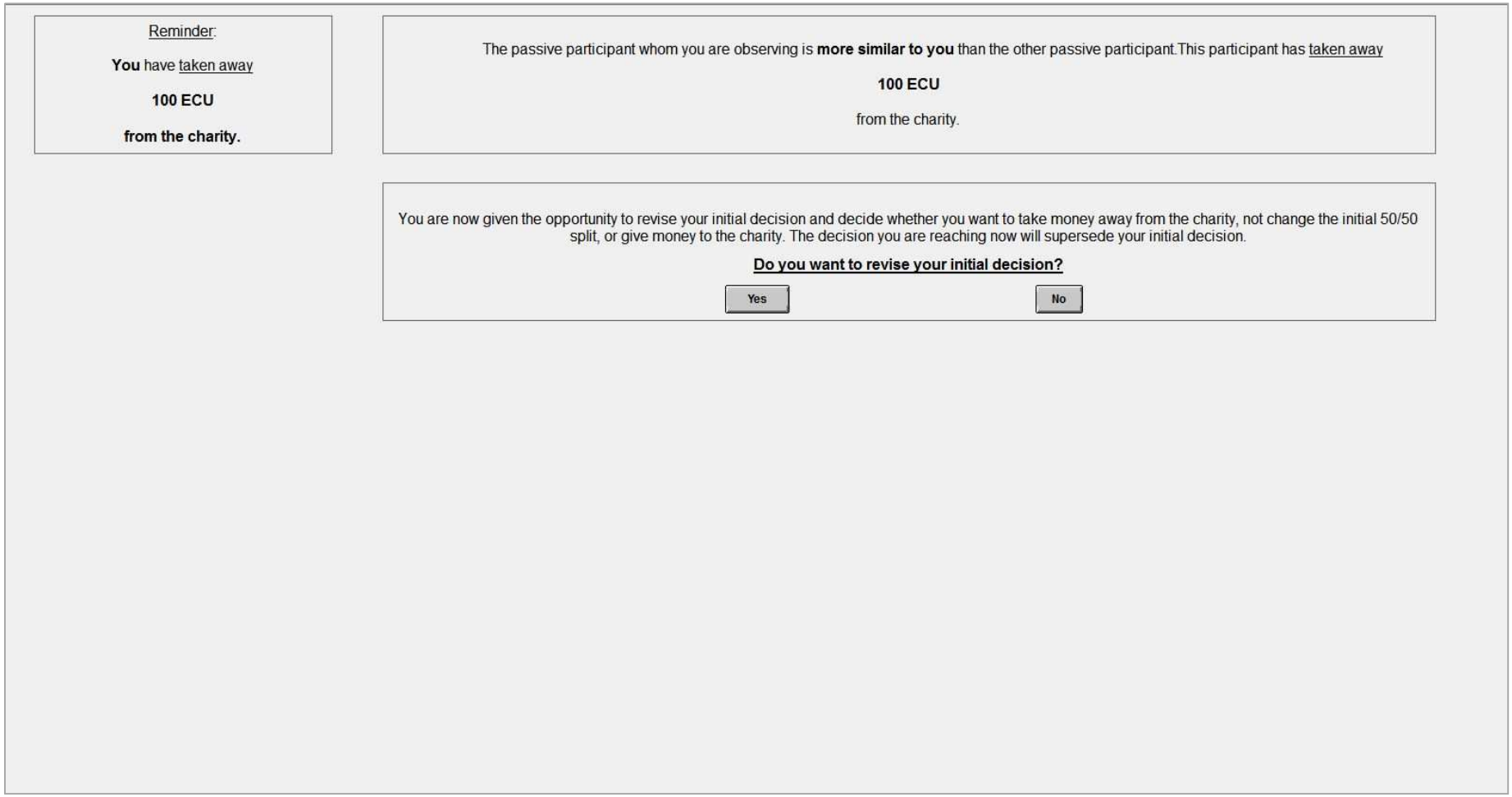


(Exemplarily for the taking away decision)



3. <u>Observation of one's peer and potential revision of one's initial decision</u>

Important: treatments vary by the information on the social proximity measure (first box): unknown similarity (Baseline), more similar (Treatment 1), less similar (Treatment 2).



(Exemplarily for revision of one's initial decision)