iscte

INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Prosocial selectivity: How behavioral information of a potential beneficiary influences decision making

João Barreiros

Doctorate in Psychology

Supervisor:
Prof. Elizabeth Claire Collins, Invited Researcher,
ISCTE – Instituto Universitário de Lisboa

Co-supervisor:
Prof. Patrícia Arriaga, Assistant Professor,
ISCTE – Instituto Universitário de Lisboa

July, 2022

Departamento de Psicologia

Prosocial selectivity: How behavioral information of a potential beneficiary influences decision making

João Barreiros

Doctorate in Psychology

Jury:
Prof. Augusta Gaspar, Assistant Professor, Universidade Católica Portuguesa
Prof. Rui Costa Lopes, Research Fellow, Instituto de Ciências Sociais – Universidade de Lisboa
Prof. Cícero Pereira, Research Professor, Instituto de Ciências Sociais – Universidade de Lisboa
Prof. Henrique Duarte, Assistant Professor, ISCTE – Instituto Universitário de Lisboa

July, 2022

Acknowledgements

Abstract

The current research examined whether young adults were less likely to help a needy target who behaved aggressively toward a third person in past, and the possible motives underlying that decision. Not helping a needy person who behaved aggressively in the past may be a form of punishment, since it may aggravate that person's need. Punishment can be administered with: a) a just deserts motive, aimed to restore justice, without the goal of changing future behavior; b) an incapacitation motive, i.e. to prevent or reduce future harms; c) a deterrence motive, i.e. to change the costs and benefits of a possible harm so it becomes unattractive. We hypothesized that participants in conditions priming a just deserts, incapacitation, and deterrence motives would volunteer less time than those in both a no-aggression condition and a punishment motives control condition. In one pretest and three experimental studies participants were asked to listen to and evaluate a bogus radio program in which a needy target is interviewed. In the aggression conditions the target reported an aggression committed to a colleague with information priming each punishment motive. Participants were then asked to volunteer time to help the target. Results of studies suggested that participants withheld help with a deterrence motive, whereas no evidence was found that confirmed the remaining hypotheses. This research shows that prosocial behavior can be withheld – when the beneficiary acted aggressively toward another person in the past – in the form of punishment with a concern for others.

Resumo

Esta investigação examinou se era menos provável que jovens adultos ajudassem um alvo em necessidade que foi agressivo para com outra pessoa no passado, e os possíveis motivos dessa decisão. Não ajudar uma pessoa em necessidade que foi agressiva no passado pode ser uma forma de punição, visto que pode agravar a sua necessidade. Uma punição pode ser administrada com os seguintes motivos: a) retributivo, visando restaurar a justiça, sem qualquer objetivo de mudar o comportamento futuro; b) de incapacitação, i.e., de prevenir ou reduzir futuras ofensas; c) de dissuasão, i.e., de mudar os custos e benefícios de uma possível ofensa tornando-a pouco atrativa.

Hipotetizámos que os participantes de condições de primação dos motivos retributivo, de incapacitação, e de dissuasão voluntariariam menos tempo do que aqueles de condições de não-agressão e de controlo de motivos de punição. Num pré-teste e três estudos experimentais pediu-se aos participantes para ouvirem e avaliarem um programa de rádio fictício no qual um alvo é entrevistado. Nas condições de agressão o alvo reporta uma agressão a uma colega contendo informação que primava cada motivo de punição. Foi depois pedido aos participantes que voluntariassem tempo para ajudar o alvo. Os resultados sugerem que os participantes inibiram a ajuda com um motivo de dissuasão, e não se confirmaram as restantes hipóteses. Esta investigação mostra que o comportamento prossocial pode ser inibido – quando o beneficiário é agressivo para com outra pessoa no passado – na forma de punição com uma preocupação por outros.

Palavras-chave: comportamento prossocial, comportamento de ajuda, comportamento agressivo, tomada de decisão, punição

PsycINFO Classification Categories and Codes:
3000 Social Psychology
3020 Group & Interpersonal Processes

Table of Contents

Index of Figures

Index of Tables

Chapter 1. Introduction

Research shows that people tend to help others that they perceive to be in need – even when escaping the situation is easy (Batson, et al., 1981; Coke, et al., 1978; Toi & Batson, 1982).

Suppose one day at the company you work for you find a new employee carrying a box full of materials toward his office, trying not to let anything fall. Surely if you could you would quickly help this person. But, if the day before you happened to witness this individual being rude to some colleagues of yours, do you think you would help so readily?

In this thesis we analyzed how the decision to help a person can be influenced by knowledge of the past behavior of that person. Regarding this relation, we reviewed previous studies from different domains:

(a) studies investigating how antisocial behavior is related to decrease helping;

(b) studies using economic games where players tend to punish non-cooperative co-players;

(c) studies involving assignment of punishment according to theories of criminal justice.

We also present studies demonstrating the influence of aggressive behaviors performed by a person on the decision to help that person, as well as the involvement of specific punishment motives in the helping behavior.

Prosocial behavior is an umbrella term that includes several behaviors that tend to benefit or increase the welfare of another person, such as helping, sharing, donating, comforting, volunteering, and cooperating (Batson et al., 2011; Dovidio & Penner, 2003; Penner, et al., 2005).

Helping has been studied by scientists for decades with many theories proposing explanations about why we help or do not, with varying degrees of empirical support (for a review see McCullough & Tabak, 2010; Penner, et al., 2005). Proposals range from egoistically motivated cognitive-based theories (e.g., The Cost-reward Analysis Model, Piliavin et al., 1969) to an altruistically motivated affective-based theory (e.g., The Empathy-Altruism Hypothesis, Batson, 2011).

On the other hand, research particularly relating prosocial with antisocial behavior is not so prolific. Vaish, Carpenter, and Tomasello (2010), for example, found that a smaller proportion of the 3-year-old children in their sample helped individuals who they had previously seen behave in an antisocial manner toward a third-party, than helped individuals who they had not previously seen act antisocially. Children also helped beneficiaries that had the intention to harm, but failed, less often than neutral beneficiaries.

In contrast, aggression is consensually defined as any behavior with the goal of harming or injuring another living being who is motivated to avoid such treatment (Anderson & Bushman, 2002). With the exception of formal, restorative justice practices (e.g., Gavrielides, 2016), no empirical research was found, to the best of our knowledge, specifically examining choices to give help to perpetrators who

are in need. However, some studies about cooperation in economic games provide some important findings in this line of research. However, those studies address non-cooperative behaviors on the part of the target of help – their actions included not giving resources to another player or not contributing to a public good (i.e., "free riding") – which differs in a number of ways from direct aggressive behavior.

Cooperation is defined by behaving interdependently with others toward a common goal that benefits all involved (Penner, et al., 2005). Fehr and Gachter (2000; 2002) showed that, in public good games (where players give resources to a common fund, which is later evenly divided by participants), non-cooperative behaviors induce strong negative emotions (i.e., anger) in cooperative individuals. Specifically, they found that the lower the contributions, the more intense were the negative emotions of those who contributed an average amount, and the more severe were the punishments enacted (consisting of deductions to the monetary units with which players were initially endowed). Negative emotions are thus correlated with punishments.

Non-cooperative behaviors may be punished with the goal of reinforcing social norms (i.e., prosocial or cooperation norms), an effective strategy for reducing the frequency of non-cooperative behaviors (Buckholtz & Marois, 2012; Carpenter, et al., 2004; Fehr & Fischbacher, 2004). Such sanctions are also used by third party observers. That is, studies have shown this effect with participants who were not involved in previous interactions with other participants (Fehr & Fischbacher, 2004). These games suggest that punishment can be performed with an egoistic or an altruistic motivation. According to Price, Cosmides, and Tooby (2002), punishment of non-cooperative individuals may have evolved as a means used by cooperative individuals to increase cooperation, and also to eliminate the fitness differential that free riders receive over cooperators, i.e., to reduce the advantage from free riding. Since cooperators potentially receive lower payoffs from a public good fund than free riders, punishing may be used to increase one's payoff. On the other hand, according to the strong reciprocity (Gintis, 2000) and the social reciprocity (Carpenter & Mathews, 2002) theories, free riders are punished (even when such action is costly) because they violate the group beneficial contribution norm, and not for accruing higher payoffs and consequently lowering the cooperators' payoffs – so in the absence of any future material gain. Fehr and Gachter (2002) also found that costly punishment provided material benefits for the future interaction partners of a punished individual, but not for the punisher.

We assume that not helping a person who is clearly perceived as being in a state of need (i.e., the potential beneficiary), but has behaved aggressively in the past, may be considered a form of punishment, since it may produce, maintain, or even increase the discomfort, dissatisfaction, suffering, or any negative affective states that result from the need that the potential beneficiary is feeling. Thus, a punishment in the form of withholding help can be used for the same reasons that drive any punishment. This research focused on criminal justice philosophies, because: (a) empirical research

2

has shown that lay people resort to punishments framed within such theories to sentence criminal cases (common sense notions of justice and punishment establish the grounds, to some extent, of those theories; see Carlsmith & Darley, 2008 and Darley & Pittman, 2003); and (b) aggressive behaviors, like crimes, are socially considered condemnable in Western cultures, besides often legally configuring crimes. Consistent with the transfer of knowledge and solving strategies that people carry from one kind of problem to another kind (Gentile, 2000), we believe those punishments can also occur in other contexts of people's lives, rather than criminal – and even assume the subtle form of withholding help. An idea that we tested in our studies, described ahead.

In psychology, research on punishment has traditionally been made in the scope of operant learning studies (e.g., use of punishments to suppress or control rates of behavior), as well as by developmental psychologists, who, in the same behaviorist perspective, addressed the use of punishments for modifying transgressive behaviors in children (Branch, 2000; Darley, et al., 2000; Weiner, et al., 1997). More interest has been shown regarding the effects of the punishment on the offender and his/her behavior rather than on the motivation of those who assign the punishment (Carlsmith, et al., 2002). Two classical broad justifications for the use of punishment have received considerable attention in psychological studies: the just deserts or retributionist philosophy and the utilitarian philosophy. These theories are described in Chapter 2.4. Punishment and negative behavioral information.

The way people use information about others when making a decision to help is rarely investigated. Why do we think some people deserve more help than others? There is still much variance in helping behavior that research has not explained. The general goal of the research in the present thesis was to understand whether young adults may withhold help toward an aggressive target, as well as the motives that might underlie that decision.

The contents of the present thesis are organized in three sections. In the following chapter, Chapter 2., we review the literature on motivational constructs and related aspects, and define egoistic and altruistic motivations. Next, we describe the theories explaining prosocial behavior, and their underlying processes, that have been developed and tested over the past decades (Chapter 3). Based on those theories and on other literature, we then explore the ways by which prosocial behavior may be withheld (Chapter 4). In this part we propose a process that may cause withholding help, which is tested in the following section, the Empirical Section – Experimental Studies. Here we describe one pretest and three subsequent experimental studies, and detail, for each study, the adopted method, the obtained results, and a discussion of how the results support or disconfirm our hypotheses. Finally, in the third section we make a General Discussion in which I reflect on the results and draw conclusions of the studies, consider their limitations, and present suggestions for future research to contribute to the understanding of helping decisions.

3

THEORETICAL SECTION: STATE OF THE ART

Chapter 2. Motivation to help

In order to understand why people withhold help from some individuals, it is certainly useful to understand why people help other individuals in the first place. The motivation processes involved in helping behaviors are not different from those occurring with other behaviors. However, since definitions of motivational constructs are not consensual in the literature, it is important to clarify the perspective adopted in the present thesis in advance. This elucidation is of particular relevance for distinguishing the altruistic from the egoistic motivations.

2.1. Need

Despite often being used in literature indiscriminately from other motivational constructs such as motive, motivational state, drive, incentive and others, *need* is used in this thesis as the lack of a resource required for the satisfaction, welfare, or survival, of an organism or individual (see Pittman & Zeigler, 2007, for a review). It is also defined by the deprivation state of an individual resulting from the lack of a resource (Reber & Reber, 2001). This means that obtaining the resource in question produces, maintains, or increases the satisfaction or welfare of the individual, while reducing or satisfying his/her need. The resource might be biological / physiological (e.g., water, food, sleep, sex), or personal / social (e.g., self-esteem, acceptance, power, social influence). As opposed to the constructs of motive, goal, and motivation that are further defined, this deficit of a resource does not have to be necessarily cognitively represented – it might only be a physiological detection (e.g., we need oxygen, but we are normally not motivated to obtain it, thus it is not our goal).

2.2. Motive/goal and motivation

Pioneer psychologists Henry Murray and Kurt Lewin conceived motives as being associated to representations of valued and desired states (Lewin, 1952; Murray, 1938). Motive and goal can be defined with the common meaning of cognitively represented states of the reality which the person believes will result from behaviors he/she will perform (Fishbach & Ferguson, 2007). In other words, we can see them as represented states to which the individual will direct behaviors (believed to produce the state).

A motive or a goal thus consists in a set of two cognitive representations: the desired state, and the belief (or expectation) that the person will behave in a way which will produce or increase the likelihood of occurrence of that state. This set might be consciously or unconsciously activated (Bargh & Chartrand, 1999).

The behaviors do not have to be necessarily specified: one might be motivated to help a neighbor getting his car fixed (or have this goal, a desired state), but not know exactly what to do – while knowing that, in order to help, he/she will have to do something.

Possible informative contents of the representations of the desired states, and multifinality

The representation of the desired state may include information about cognitions, affects, and/or the individual's or other people's behaviors. For instance, I might want to help my neighbor because I want to: know more about car mechanics (cognition); feel less distressed (affect); feel good (affect); or maybe because I expect him to help me cleaning my garden next week (behavior). And of course, I might believe the same behavior will produce multiple states of reality (e.g., I will know more and feel good), which researchers called multifinality (Dunning, 2011).

Positive hedonic effect

These desired states (whose representation is activated when activating a motive) are assumed to be valued or positively evaluated by the individual (Fishbach & Ferguson, 2007; Gollwitzer & Brandstatter, 1996), which means that he/she believes that its attainment will provide pleasure, satisfaction, or another positive affect. This also happens with particular represented states which are not expected to be pleasant themselves, but to cause desired states – these ones being pleasant; though unpleasant, the former represented states are also valued by the individual because he/she believes they will cause pleasant, desired states (see Instrumental/proximal motives, and ultimate/distal motives, ahead).

We can also assume that such hedonic effect indeed occurs, i.e., experiencing a state for which the individual has dedicated behaviors produces positive affect.

Anticipated and non-anticipated consequences

However, the individual does not have to act necessarily in order to experience such positive affect (at least not consciously), even though believing or knowing that that will happen. See for instance that I might help my neighbor knowing that I will enjoy having contributed to getting his car fixed, but I actually did it because I wanted to enlarge my knowledge of mechanics. This means that, in principle, if I was somehow incapable of extracting any satisfaction from having the car fixed – knowing it in advance or not –, I would help anyway, because my motive was to enlarge my knowledge, not to extract satisfaction.

In sum, the individual might believe that reaching a state of affairs will be satisfactory (which indeed will likely happen) – but such positive affect does not necessarily have to be what motivated him/her.

Frequently a certain behavior produces several states, with only some of them having been activated as representations – i.e., only some of the occurred states, of all that occurred, may have been motives of the behavior which produced them. The produced states which were not previously activated as representations by the subject – which were not motives of his/her behavior – are called unanticipated states, or unintended consequences (Batson, et al., 2011; Batson, 1994). Of all the anticipated states that the individual may activate, some may not, of course, be desired; the individual knows that attaining a certain state will produce, in turn, another state that is not positively valued – but which will also not prevent him from acting anyway.

This issue is particularly important for understanding the altruistic motive, which is discussed in Chapter 3.2. The theory of prosocial behavior with an altruistic motive – The Empathy-Altruism Hypothesis.

Instrumental/proximal motives, and ultimate/distal motives

Another important notion about motives is that they can be causally related by the individual. In fact, making a decision can illustrate our knowledge about how events relate in the world, and how we take advantage of that. People know that sometimes in order to achieve one desired state it is necessary to achieve a previous one, which will cause the desired state. For example, I might be motivated to fix (behavior) my car (first state) so that I can travel to my country house on the weekend (desired state). Or I can help (behavior) my neighbor fixing his car (first state) expecting that he will help me cleaning my garden next week (desired state). The first states are thus believed or expected: (a) to result from behaviors that the individual plans to perform; and (b) to cause the desired states, for which they are called instrumental or proximal motives/goals. The desired states are called ultimate or distal motives/goals (Batson, et al., 2011; Park & Buunk, 2011).

Motivation and willingness

So far I have been referring to motives or goals, meaning represented states that the person believes will result from behaviors. However, sometimes the expressions "having the motivation" or "willingness" to do something, or "being motivated to" are used. In these cases, motivation has the simple sense of believing that one will likely perform certain behaviors. For example, when I say that "I am motivated to help my neighbor" it means that I believe I will likely perform helping behaviors. Or when I say that "I was motivated to fix my car, but I when I tried, I lost my motivation", it means that I ceased to believe that I would likely perform the behaviors.

Intention shares this definition, but is generally considered a conscious belief, or a conscious process of decision-making (Reber & Reber, 2001; VandenBos, 2015).

It is important to consider some relevant motivational factors that might influence behavior, beyond the activation of needs and motives.

First, the probability of the occurrence of a behavior depends on the strength or intensity of the motive that compels that behavior (Batson, et al., 2011). It is possible that this property is:

(a) the probability estimated by the individual for the occurrence of the behavior involved (e.g., if my motivation for fixing my car is not strong, it is because I do not believe it is likely that I will do it);

(b) the probability estimated by the individual that the involved behavior will produce the desired state (e.g., if my motivation for having the car fixed is low, it is because I do not consider it likely that my intervention on the engine will fix the problem);

(c) the frequency of activation of the belief, either conscious or unconscious (e.g., if my motivation for fixing my car is low, it is because I do not think of it very often).

Another motivational factor influencing behavior is the activation of other motives which are incompatible, or that make the attainment of the motive in question impractical. For instance, I might want to help my neighbor fixing his car, but I also need to go to work, and I do not have time for both. The motive of going to work thus makes it less likely that I will help my neighbor (Fishbach & Ferguson, 2007).

Finally, the activation of representations of other behaviors believed to produce the same desired state might also compromise a specific behavior (Batson, et al., 2011). For example, I want my car fixed, and I can try to do it by working on the engine, but I can also ask my brother to do it. This can reduce my motivation to work on the engine.

Chapter 3. Theories of prosocial behavior

3.1. Common assumptions across prosocial behavior theories: A psychological process

Theories of prosocial behavior assume that the benefactor's prosocial behavior results from a psychological process with two fundamental stages: the activation of information about potential needs of beneficiaries, and the decision-making process regarding (in)action, which involves believing that some particular behavior will satisfy the potential beneficiary's need. In effect, independent of other variables that may intervene in this second stage (such as the activation of information about a potential benefactor's particular own need, like in the reciprocity theories), theories assume that the potential benefactor believes his/her behavior will produce the specific effect – i.e., an anticipated state – of promoting the satisfaction, reduction, or suppression of a potential beneficiary's need, which in turn will increase the potential beneficiary's welfare. The potential benefactor knows that the potential beneficiary lacks some resource affecting his welfare, which the potential benefactor can provide.

The instrumental motive – satisfying the potential beneficiary's need

The activation of information about the potential beneficiary's need may result from perceiving it, or inferring from observing his/her behaviors – when the potential benefactor has some sort of contact with the potential beneficiary –, but may also be known, when somebody tells the potential benefactor about the need, or when he/she learns about the need in a document or other media (e.g., when a doctor first reads the clinical file of a patient). The potential benefactor may also believe that the potential beneficiary will experience the need in the future, or expect it, like for instance when the potential benefactor knows that a relative died, and expects that his son will be very sad when he learns about that fact, and thus will be in need of comforting.

Prosocial behaviors are performed with the instrumental or proximal motive of promoting the satisfaction of a potential beneficiary's need. The ultimate or distal motive may be to produce, maintain, or increase the potential beneficiary's welfare (Batson, et al., 2011; Batson & Powell, 2003; Dovidio & Penner, 2003), the potential benefactor's own welfare (Batson, et al., 2011; Dovidio & Penner, 2003), or even the welfare of both (Batson & Powell, 2003).

The ultimate motive – providing welfare to a target beneficiary

When the potential benefactor acts with an ultimate motive of increasing his own welfare, while satisfying another's need, then it is considered to be an egoistically motivated prosocial behavior. On the other hand, if the potential benefactor wishes to increase the welfare of another person without concern regarding obtaining any personal reward or benefit, then he is altruistically motivated (Krebs,

9

1982). One important debate in the study of prosocial behavior focuses precisely in the distinction between the egotistically and the altruistically motivated prosocial behavior (see 3.4. The egoism/altruism debate).

The cognitive process of behaving prosocially with an egoistic motivation

We can now understand that the anticipated state of increasing a potential beneficiary's welfare may not be the only anticipated state that the potential benefactor believes will accrue from his behavior. Other outcomes involving the potential benefactor himself may be expected, such us experiencing feelings, fulfilling norms, and inducing behaviors in others – all resulting in his own increase of welfare (see Figure 3).

Figure 3

A Potential Benefactor's Cognitive Process of Behaving Prosocially with an Egoistic Motivation
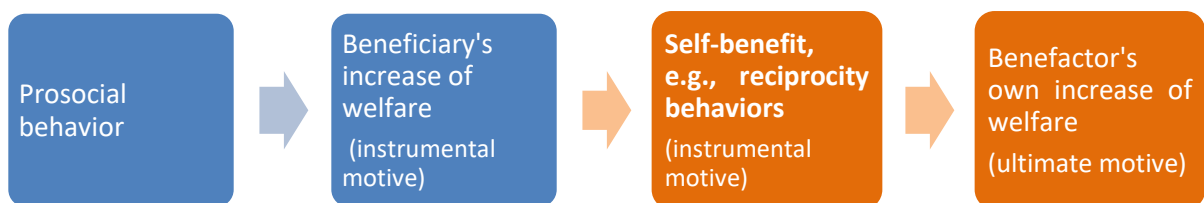
An egoistic motivated prosocial behavior thus implies that the potential benefactor believes or expects: (a) his behavior will satisfy another person's need and increase this person's welfare – the instrumental motive; (b) such increase in welfare will cause, in turn, some self-benefit – the ultimate motive.

Different types of self-benefits include obtaining rewards (e.g., feeling better or receiving money), avoiding punishments, benefiting from others' behaviors, and are described in the following chapters.

3.2. Theories of prosocial behavior with egoistic motives

3.2.1. Characterization of self-benefits

Self-benefits can be understood as resources – namely material, affective states, behaviors of others, or events, like avoiding aversive situations – that produce or increase welfare. As explained earlier, the obtainment of a resource may be anticipated by the individual as a desired state (the state of

experiencing the resource or the event, believed to cause welfare), which he/she believes will result from prosocial behaviors, thus constituting a motive.

Batson and his colleagues (Batson et al., 2002; Batson et al., 2008) identified different types of self-benefits resulting from prosocial behaviors, which can be organized into three categories: aversive situation or punishment avoidance self-benefits, aversive arousal reduction self-benefits, obtainment of resources.

Self-benefits that fall under the label of avoidance of aversive situations or punishments are self-benefits whose activation as motives imply that the individual, before acting prosocially, imagines two courses of events or states: one where he/she does not behave prosocially, causing an aversive situation (e.g., feeling guilt), and another where the individual does behave prosocially and thus prevents that aversive situation – which is preferred and enacted. The potential benefactor believes that helping another person will prevent him from receiving some punishment or experiencing some negative affective state (and thus reducing his/her welfare). For example, when seeing a passerby stumbling on something and falling in the street, the potential benefactor may be compelled to help only to avoid the censure of those around him/herself, that he knows would occur if he/she withheld the help. Sometimes it is no longer possible to avoid an aversive situation. When already experiencing some negative affective state (e.g., due to seeing a person suffering), the self-benefit desired is to feel better or to reduce the aversive arousal. Some theories propose psychological processes that people activate with this ultimate motive, labeled as the reduction of aversive arousal or negative affective state self-benefit. Finally, there are other self-benefits that do not require either anticipating two courses of events or experiencing any negative affective state. They are just resources that individuals believe they will obtain after behaving prosocially. They are in the category of obtainment of resources.

Self-benefits also differ in their nature. Whether they are resources believed to be obtained, or aversive situations believed to be avoided with some prosocial behavior, they can be material (e.g., receiving a payment; avoiding a fine), social (e.g., thanks; censure), affective states (e.g., empathic joy; guilt), or cognitive (e.g., fulfilling a personal or social norm; failing to attend to a norm).

Table 1 describes the self-benefits that can result from acting prosocially toward another person. In this table, the reduction of aversive arousal or negative affective state theories were included in the avoidance of aversive situations or punishments category.

Table 3

Possible Self-benefits Resulting from Behaving Prosocially (based on Batson et al., 2008)

Obtainment of resources			
Cognitive:	**Affective:**	**Social – behavioral:**	**Social – material:**
● Personal norms ● Social norms	● Mood enhancement or maintenance ● Empathic joy ● Self-esteem ● Pride ● Honor	● Gratitude ● Praise ● Esteem or social approval ● Reciprocity	● Receive a payment ● Gifts
Avoidance of aversive situations or punishments			
Cognitive	**Affective:**	**Social – behavioral:**	**Social – material:**
● Cognitive discrepancy ● Injustice	● Arousal reduction ● Relief ● Guilt ● Shame	● Attack ● Censure ● Sanctions for norm violation	● Fines

3.2.2. Obtainment of resources

Cognitive resources

Schwartz (1977) suggested that people can behave prosocially in order to respect personal norms: rules that indicate how to behave in specific situations, having a sense of obligation, which are socially learned and vary between individuals of the same society. According to this theory, to act prosocially one does not need to experience any emotion when perceiving someone in need, or to expect any reciprocal behavior from the potential beneficiary, but only needs to have the idea that some personal norm will be respected when behaving prosocially.

The impact of personal norms on prosocial behaviors seems to depend on how much attention is paid to such norms in particular situations. Focusing the attention inward can lead one to compare oneself to some important standard or ideal behavior (Gibbons, 1990). If these standards involve helping, then a person is more likely to help (Hoover, et al., 1983).

People use personal and social norms to guide what they should do and how (Higgins & Spiegel, 2004). When people do not comply with such standards, they can feel bad (Williamson, et al., 1996). Satisfying a personal norm may then be rewarding in itself, but also a way of avoiding negative feelings.

Similarly, principles – rules or standards of behavior considered moral (VandenBos, 2015) – can also motivate prosocial behavior. Examples are the principle of fairness or justice, or the utilitarian principle of the greatest good for the greatest number. Currently, the empirical relation between moral principles and prosocial behavior is weak (Batson, 2011). This can be due to the fact that principles lack a strong emotional base, as well as to a tendency that people have to defend principles, but fail to act according to them (Blasi, 1980). In any case, some moral principles, such as the principle of universal justice (Kohlberg, 1976), are associated with increased prosocial behavior (Eisenberg, 1991; Emler, et al., 1983; Erkut, et al., 1981; Sparks & Durkin, 1987).

According to the principle of care, one should help those in need (Batson, 2011; Hoffman, 2000; Eisenberg, 1986). Wilhelm and Bekkers (2010) found that the principle of care was strongly associated with ten types of helping behavior. A few years later Bekkers and Ottoni-Wilhelm (2016) also found a positive association between the principle of care and charitable giving to people in need.

Prosocial behaviors can also be used to promote personal values. Values are states of reality that are desired, for which people strive; in this sense, they are goals or motives, but with the particularity of being trans-situational, i.e., the same behaviors and states can occur in different situations or contexts. Because of this amplitude of application, they are abstract and general (e.g., freedom, equality, friendship, courage), and benefit the community, not only the individual (Malle & Dickert, 2007; Schwartz, 1996). As standards of the desirable, values guide evaluations, choices, and behaviors (Prentice, 2000).

Lonnqvist and colleagues (2013) found that personal values, particularly universalism (a value that includes protection for the welfare of all people) predicted cooperation in a prisoner's dilemma interaction. In this game two participants who possess an initial quantity of resources (e.g., monetary units) are given the choice, simultaneously, to provide any amount of those resources to each other (cooperate) or not to give any at all (defect; in the original version decisions to give resources are made in turn). The players know in advance that any amount they provide to each other is doubled, and that after they both make their decisions, they will receive a reward, whose value depends on their decisions: they receive a more valuable reward if they both cooperate than if both defect, and if one player cooperates and the other defects, the cooperator receives the smallest payoff, whereas the defector the most valuable reward.

In a one-round version of the game, defection provides the best outcome, regardless of the choice of the coplayer. But with multiple rounds, players start taking into consideration the previous decisions of their coplayers in their own decisions. Noticeably, players tend to respond to their coplayers in the

13

same fashion: if one defects then the other defects too, and if one behaves prosocially, the other will reciprocate (the so-called "tit-for-tat" strategy; Axelrod, 1980). If players have been defecting, and one starts or resumes cooperating, the other one will cooperate immediately, suggesting a forgiving gesture.

People can also behave prosocially with a motive of respecting what is considered socially appropriate, desired, and expected behavior. Social norms are rules of appropriate behaviors that people learn by directly being rewarded or punished for what they do, and by social learning (Gürerk, et al., 2006; Weber & Murnighan, 2008). These rules are shared by individuals within groups (Schwartz & Howard, 1982).

As people learn that behaving prosocially can bring rewards, they learn the norms involving prosocial behaviors that should be performed in different social situations. Some norms prescribe that one should help another who is in need, in certain circumstances, to avoid self-administered or social sanctions (Batson & Powell, 2003).

Different social norms predict and are associated to prosocial behavior. Three social norms frequently studied in the context of prosocial behavior are the social responsibility norm, the fairness or equity norm, and the reciprocity norm. According to the social responsibility norm it is expected that people help anyone who is dependent on them (Berkowitz, 1972; Dovidio, 1984). Berkowitz and Daniels conducted an experiment in 1963 where they asked students to assemble envelopes under the supervision of another student, telling them that the work would be a test of the supervisor's skills (they would then be helping someone dependent on them). The researchers found that students were more motivated to work when the consequences for their supervisor were greater, even when there was no possible gain to the students themselves.

There are, however, situations where the potential benefactor might feel free from respecting the norm, or even inhibited from acting due to the activation of other conflicting norms. For example, in a supermarket, if a child is trying to reach a bag of candies placed on a shelf that is too high, and his father noticed the situation and did nothing, the potential benefactor may be compelled to give the bag to the child – proceeding accordingly to the social responsibility norm –, or be inhibited by a norm of respect for the family's own liberty of choice, or by a norm of respect of the parents' authority. The potential benefactor can also be less willing to follow the norm if the interdependence relationship with the needy person is not desired, if it threatens the potential benefactor's feelings of personal freedom, or if the person in need is perceived as responsible for his own state due to lack of effort or immoral conduct (Berkowitz, 1973; Frey & Gaertner, 1986; Weiner, et al., 1988). This last case is related to the research completed for this thesis.

When the potential beneficiary is perceived as responsible for his/her own state of need

Irving Piliavin and colleagues (1969), for example, showed that an apparently ill person was more likely to receive aid than another who appeared to be drunk (both collapsing on the floor inside a subway car). Likewise, Bernard Weiner (1980) found that judges considered that students would more likely lend class notes to unknown classmates when the need of notes was attributed to uncontrollable factors (e.g., classmates' ability to take good notes, or teacher shortcomings) than when it was attributed to controllable factors (e.g., classmates' lack of effort). Another, more recent example comes from the clinical context: Michelle Lobchuk and colleagues (2008) found that caregivers of lung cancer patients gave less supportive help if they believed that the patients were largely responsible for their disease.

Another social norm driving prosocial behavior is the equity or fairness norm. The equity theory affirms that people activate motives based on norms to preserve and restore equity, defined as a state in which the ratio between what is given and received (like costs and benefits) is equal for all the individuals in a relationship (Walster, et al., 1973). People have, in fact, an aversion to unequal treatment. For example, in prisoner's dilemmas people are much less cooperative when their partners systematically receive better payoffs than they receive (Sheposh & Gallo, 1973). Similarly, when people feel they have been undercompensated for their work, they are less helpful to their colleagues (Organ & Ryan, 1995).

Moreover, people mobilize efforts to avoid inequity – namely through prosocial behavior. Marwell and Schmitt (1972) showed in an experiment that participants who unfairly received too much monetary reward based on their contributions to the group activity often freely chose to forego some of that reward to avoid inequitable conditions. Plus, when they had the opportunity to rectify the inequity, most participants transferred sufficient amounts of rewards to their peers to produce equity.

Finally, the norm of reciprocity states that we should benefit, refrain from harming, or relieve from pain, people from whom we received benefits in the past. It also adds that we should not help those who denied helping us in the past without a legitimate reason (Gouldner, 1960). The pressure to comply with the norm depends on the past circumstances in which the help was given to the subject in the first place, including: (a) the intensity of his need; (b) the perception the subject has of the motive that led the other person to help him in first place; and (c) the perception the subject has of the total amount of help given by the other person, relatively to his resources (Gouldner, 1960).

Consistent with this theory, Boster, Fediuk and Kotowski (2001) demonstrated that participants were more likely to help other participants who offered help before. Pruitt (1968) also showed that the amount of help given by participants after receiving help was a direct function of the amount of help previously received, suggesting that people are motivated to "pay back", or "compensate" the help they are given. A few years later Kahn and Tice (1973) corroborated this finding.

15

Affective resources

Research shows that people are more likely to help when they are in a good mood (Baron, 1997; Isen & Levin, 1972; Salovey, et al., 1991; Weyant, 1978). Participants who already are in a good mood, experimentally induced, tend to help a person in need more than other participants whom were not in an induced good mood (Isen & Levin, 1972; Isen & Simmonds, 1978). Participants in a good mood may help more because they believe that such help will maintain their mood.

After witnessing someone in need of help, people tend to experience a negative affective state, which they try to reduce or relieve by means of helping that person (or other means). Prosocial behavior performed in this particular situation, to increase mood from a negative level (Cialdini, et al., 1973) is described ahead (see Negative state relief).

Helping may also promote a good mood (Midlarsky, 1991; Williamson & Clark, 1989; Yinon & Landau, 1987). However, these studies only demonstrated that the help provided by participants *resulted* in improved moods – not that such emotional states necessarily constituted the motives for helping (see Chapter 2.2.). There is, nonetheless, the belief that helping enhances the mood. For example, in responding to a survey, students reported believing that donating money, helping to pick up papers, or giving a ride to a hitchhiker would improve their moods (Harris, 1997).

Another affective resource that may motivate helping is empathic joy. Empathic joy is the pleasure or joy felt by a benefactor at seeing a beneficiary experiencing relief from some need (Hoffman, 1981). Smith, Keating, and Stotland (1989) found that participants empathically aroused (for empathy and empathic concern see 3.3. The theory of prosocial behavior with an altruistic motive – The Empathy-Altruism Hypothesis) were more likely to help a target when they believed that they would receive feedback on the results of their help – i.e., feedback conveying that the beneficiary's need had been reduced or satisfied – than when they believed they would not receive feedback. Feedback provided by the potential beneficiary was thus shown to be essential for the helper to experience empathic joy. Consistently, in subsequent research conducted by Batson and colleagues (1991), participants who experienced low levels of empathic concern were more likely to help when they believed they would learn the benefits of their intervention (i.e., feedback), particularly when that intervention would likely improve the potential beneficiary's situation – thus allow them to feel empathic joy. Interestingly, expecting feedback significantly increased helping among participants induced to feel low empathy, but not among those induced to feel high empathy.

People may also help to enhance their self-esteem. Self-esteem is defined as the feelings produced by one's self-concept (i.e., cognitive representation of oneself; Leary, 2003; Sedikides, 1993). Global or general self-esteem – the feelings produced by the self-concept as a whole – is distinguished from specific self-esteem, which refers to the feelings produced by a particular trait or characteristic of the self-concept (e.g., physical appearance). These feelings can vary over time (Crocker & Bylsma, 1996).

16

People seem to have a need to maintain and increase their self-esteem (Leary & Downs, 1995), and a positive self-representation (Crocker & Bylsma, 1996). Self-esteem may promote one's well-being because, as Taylor and Brown (1988) suggest, the feelings of control typically associated to a high self-esteem encourage people to make effort to reach their goals and face adversities.

People can behave prosocially with a motive of increasing or restoring their self-esteem (Cunningham, et al., 1980; Regan, et al., 1972). Brown and Smart (1991) measured the self-esteem of students and gave them feedback of success or failure on an alleged test of their intellectual ability (having manipulated the difficulty of the test). They found that those students classified as having a high self-esteem offered more time to help a target than those with a low self-esteem – but only after being given failure feedback. This result suggests that high self-esteem students helped in order to increase and restore their damaged self-esteem. In a similar replication of this study, Yagi and Shimizu (1996) also found that their participants helped more when they received fictitious poor results of a bogus personality and intelligence test than when they did not receive any feedback.

An individual may also behave prosocially with a motive of enhancing the esteem that the potential beneficiary, or others who perceive or come to know about the helping behavior, have for him. Such motive has also been named social value, social acceptance, and social approval (hereafter designated only by esteem; Leary & Allen, 2010). Like self-esteem, esteem refers to the feelings generated by a cognitive representation of a person that are experienced by other people (both the representation and its associated feelings; Leary, 2003). So, an individual acts, namely prosocially, to produce an image of himself, hopefully agreeable, in other people. It is not surprising that we have this motivation, if we think how important it can be for cultivating sound relationships and obtaining favorable outcomes from them.

Hardy and van Vugt (2006) conducted a public good game in which participants were initially endowed with a certain amount of monetary units and were asked to contribute with any amount to a group fund and/or to a personal fund. They did not know how many rounds they would play. In a reputation condition, at the end of the first-round participants received feedback of the contributions made by all the members of their group, whereas in a control condition participants did not. Participants in the reputation condition contributed more for the group fund in a second round played, i.e., after knowing that the amount of their contributions could be disclosed to the members of their group, than in the control condition, suggesting that they behaved prosocially to increase their reputation.

Clary et al. (1998) created an instrument to measure the motivations underlying volunteering, and found that among other motives, respondents reported that they volunteered to earn the approval of their friends.

Thus, with respect to esteem, people can behave prosocially with the following motives:

- To produce a positive cognitive self-representation, or self-concept (i.e., to enhance self-esteem);
- To produce in the beneficiary (or another person who knows about the behavior) a positive cognitive representation of (i.e., to enhance social approval).

Social and behavioral resources – reciprocity behaviors

Originally termed reciprocal altruism (Trivers, 1971), the direct reciprocity theory is based on the principle that an individual may behave prosocially if two conditions are met: (a) the probability that a benefactor and a beneficiary will meet again in the future is higher than zero; and (b) the probability that a beneficiary will be available to help his benefactor, if needed in the future, is also higher than zero (Nowak, 2006). In other words, a potential benefactor may behave prosocially if he believes that after helping the potential beneficiary, he will meet him again in the future, and will be available to receive his help, should the need arise. Thus, a direct reciprocity relationship is established when a beneficiary behaves prosocially toward his previous benefactor (or repays the prosocial behavior).

In this perspective, the retribution expected by the potential benefactor, when behaving prosocially toward another person, can range from a simple "thank you" to a raise of a salary, after helping one's boss picking up his dropped papers.

One example of a direct reciprocity motivated behavior is given by waiters of restaurants. Rind and Strohmetz (1999) showed that when a server wrote a helpful message about an upcoming dinner special on the backs of the checks (i.e., alerting for a good deal), before delivering them to the customers, the size of tips given to that server increased.

Direct reciprocity has been extensively studied in the context of economic games, such as the prisoner's dilemma (Rapoport & Chammah, 1965; see Cognitive resources).

Another prosocial interaction that can be anticipated by the potential benefactor, which does not require that he meets the beneficiary again after a first interaction, is indirect reciprocity. In this particular interaction the retribution does not come from the beneficiary of a prosocial behavior, but from other people who perceived or learned about such behavior, acting due to the benefactor's good reputation (Nowak, 2006). Indeed, people tend to help and cooperate more frequently with individuals who helped and cooperated with others in the past, rather with individuals who did not behave like that (Seinen & Schram, 2006; Wedekind & Milinski, 2000). Thus, prosocial behavior creates a good reputation, and increases the likelihood that a third person will act prosocially toward him (Nowak & Sigmund, 2005; 1998). Helping someone with a good reputation may occur because these people are more likely to reciprocate later than someone with a bad or unknown reputation.

Wedekind and Braithwaite (2002) empirically demonstrated that offering help increases an individual's reputation among the members of a community.

Simpson and Willer (2008) found that participants *a priori* classified as egoists acted prosocially when there were opportunities for indirect reciprocity and reputational incentives, but not in the absence of these conditions. Participants classified as altruists, on the other hand, were less affected by reputational incentives, and tended to act prosocially regardless of the presence of reputational incentives.

These interactions have been shown with children too. In studies using a third-person giving task (Olson & Spelke, 2008), children were introduced to a protagonist doll and helped her to allocate resources (e.g., miniature plastic bananas, oranges, seashells) to other doll actors. Before making these helping decisions, children read brief scripts describing the actors' previous activities. In one of these studies, in which children read a script involving actor dolls (givers) who gave resources to other actor dolls, and actor dolls who did not perform such behaviors (non-givers), children led the protagonist doll to give more resources to givers than to non-givers.

3.2.3. Avoidance of aversive situations or punishments

Cognitive aversive situations or punishments

Janusz Reykowski (1982) proposed that if an individual perceives a discrepancy between a current state of affairs – e.g., another person in a state of need – and an expected or ideal state – the person without the need –, he will be motivated to reduce that discrepancy, in this case via behaving prosocially toward the needy person. Thus, it is the inconsistency between what is perceived and desired, or expected, that triggers the individual's motivation and consequent behavior (to reduce that inconsistency).

It should be noted that although the activation of a discrepancy between a perceived current state of affairs and another cognitively represented state, valued and desired, is needed for producing any behavior that aims to attain that state (see 2.2. Motive/goal, and motivation), the ultimate motive according to this theory would be to reduce the referred discrepancy in itself – not to increase the beneficiary's welfare, or the benefactor's, or to get other self-benefits described in previous sections.

If an individual believes that people get what they deserve and deserve what they get – the just-world hypothesis (Furnham, 2003; Lerner, 1980) – then perceiving a person in need, believed to be innocent, threatens this belief. Like with the cognitive discrepancy reduction process, such inconsistency between what one experiences and what one believes produces the motivation to reduce it, by helping the needy person. Once the need is reduced the just-world belief ceases to be

threatened. Thus, people who hold the just-world belief may behave prosocially to protect or restore their belief.

Consistent with this prediction, Foster and Rusbult (1999) found that people with stronger beliefs in a just-world were more motivated to achieve positions of power that allowed them to share resources to compensate others for perceived injustices.

According to the just-world belief, thus, not everyone deserves to be helped. Linda Skitka (1999) found that people are more likely to support federal disaster assistance (in the U.S.) to communities that take precautions against flood damage than to those who do not take such precautions. In a study conducted by Hanna Zagefka and colleagues a decade later (2010), they found that naturally caused disasters elicited more donations than humanly caused disasters, due to the participants' perception that victims of natural disasters are to be blamed less for their plight, and make more effort to help themselves.

Mary Turner DePalma and colleagues (1999) provide another example from a medical context: when given an actual opportunity to help a medical patient, participants were more likely to help an individual presented as not responsible for his/her disease onset (vs. responsible) and helping was higher in participants who also held a high belief in a just world (DePalma, et al., 1999).

In a previous study conducted with colleagues, Skitka also showed that conservatives (vs. liberals) withheld assistance from people responsible for their predicament in a variety of policy domains (Skitka & Tetlock, 1992). These researchers explored explanations for this result in a further study (Skitka & Tetlock, 1993), and found that conservatives were motivated to punish violators of social norms and to deter free riders. We addressed these motivations for withholding help in the research we conducted ourselves (see Chapter 5).

Affective aversive situations or punishments

Perceiving a person in a state of need may produce an emotion in the perceiver. Such arousal may trigger, in turn, a prosocial motivation to reduce or extinguish the person's need. Many theories have sought to explain this process and some obtained large empirical support (see, e.g., Eisenberg & Miller, 1987 for a review). One first difference that can be established between these theories lies in the ultimate motive that the potential benefactor activates subsequent to perceiving the need and experiencing the arousal: if the ultimate motive is to reduce his own aversive arousal or distress (or to increase his own welfare) then he is egoistically motivated – as the aversive arousal reduction (Piliavin et al., 1982) and the negative state relief (Cialdini et al., 1987) theories imply –, whereas if the ultimate motive is to increase the welfare of the needy person, then he is altruistically motivated – as the empathy-altruism hypothesis proposes.

Another important theory based on affective arousal is the cost-reward analysis model (Piliavin, et al., 1969). This model proposes that the occurrence of a prosocial behavior may also depend, beyond the affective arousal, on a rational and economic decision process (see 3.2.4. The Cost-reward analysis model).

According to the aversive arousal reduction theory (Piliavin, et al., 1982; Hoffman, 1981; Dovidio, et al., 1991; Schroeder, et al., 1995), the potential benefactor who perceives another person in need, particularly in an emergency situation, becomes upset or distressed, and consequently motivated to eliminate the stimulus (or the cause) of that aversive arousal. One way of reducing or suppressing such stimulus is by helping the person in need. Thus, relieving the other person's need is an instrumental motive to the ultimate motive of feeling better.

Similar to the aversive-arousal theory, the negative state relief model also suggests that perceiving another person in need produces in the potential benefactor an unpleasant, negative affective state, which in turn will activate the motivation to help. However, in this case the potential benefactor is not motivated to eliminate the cause of his distress, but rather to obtain resources or rewards that he learned were associated with prosocial behaviors. So, just like with the aversive-arousal theory, the potential benefactor is motivated to increase his welfare, but satisfying the beneficiary's need will not be sufficient to such increase. Only the resources or rewards that result from helping (e.g., social approval) will make the potential benefactor feel better.

Cialdini, Darby, and Vincent (1973) showed that people who felt bad because they accidentally harmed someone (ruining a student's thesis data), or saw another person harming someone (witnessing another person ruin the data), were more likely to volunteer to make phone calls for a worthy cause than people who were not induced to feel bad. Yet, when people induced to feel bad had their negative state relieved by receiving praise or a dollar before having the chance to volunteer, they no longer helped.

Consistently, anticipating that one's mood will be improved may reduce the motivation to help. Schaller and Cialdini (1988) showed that people who felt sad helped less when they expected that they would soon listen to a comedy tape than when they thought the tape would not be funny.

Despite its logical plausibility, and the reception of some empirical support, the motivation this theory proposes has been questioned. As Batson, Ahmad, and Stocks (2011) put it,

> Although the negative-state relief explanation received some initial support (Cialdini et al., 1987; Schaller & Cialdini, 1988), subsequent research revealed that this support was probably due to procedural artifacts (e.g., distraction). Experiments avoiding these artifacts have consistently supported the empathy–altruism hypothesis (Batson et al., 1989; Dovidio et al.,

1990; Schroeder, et al., 1988). It now seems clear that the motivation produced by empathic concern is not directed toward the egoistic goal of negative-state relief. (p. 112)

Sanctions for norm violations

We saw in Chapter 3.2. that people may act prosocially to attend to social norms – specifically, helping someone who is dependent on the benefactor (social responsibility norm), helping another to restore the balance between what is given and received in a relationship (equity norm), or help those from whom we received previous help ourselves (reciprocity norm).

But it is also true that people may help in similar circumstances not exactly because they are concerned to do the right thing according to these norms, but rather only to avoid what would happen if they did not comply with them, i.e., avoid a sanction or punishment. Such sanction can be administered by other people, self-administered, or both.

For example, when I decide to help an old lady to cross the street, I may do it in response to the social responsibility norm. However, I might instead help in order to avoid the censure of other people who are approaching the crossing and see us, and/or shame. Some studies have found that people sometimes may indeed help out of reputational concerns. For example, Piazza and Bering (2008) showed that their participants were more generous in donating to a target when they knew that this target would later discuss the amount donated with a third party, who knew the participants, than when there would be no discussion with another person.

Here is an example of helping with an apparent reciprocity norm motivation: in a work meeting I am assigned a duty to accomplish in a too short time, and a colleague of mine, present in that meeting, makes an offer of help, which I accept. A couple of weeks later, in another meeting, it is this colleague who is, in turn, in need of help. Assuming I am capable of and available to help, I offer to help, respecting the norm of reciprocity; or to prevent feeling guilt or regret (and/or the negative regard of the others who were present in both meetings). Experimental games show that failing to comply with the norm of reciprocity may even trigger punishment behaviors. Participants punish those who do not reciprocate, even when doing so is costly to them (altruistic punishment; Fehr & Gächter, 2002).

Everyday situations involving prosocial behaviors apparently motivated by equity concerns may be subtle and less common. Sometimes it happens, for example, that when we are shopping in a grocery store and paying, the seller gives us too much change, by mistake. Facing this situation, we decline the extra and explain the mistake (in line with the findings of, e.g., Marwell and Schmitt, 1972; see in Chapter 3.2. Cognitive resources). We may do it out of a fairness concern, or to avoid feeling guilty later. Reis and Gruzen (1976) conducted a study in which each participant had to distribute a certain payment between himself and three other partners. Results showed that when participants

22

had to inform the experimenter of their decisions, payment distributions were more equitable. Likewise, when participants' peers were aware of the allocations, equality-based solutions were more favored. Also in experimental games, participants who do not comply with egalitarian distribution norms are even punished by other players who are unaffected by these norm violations (Fehr & Fischbacher, 2004). These effects possibly reflect what happens in real life.

3.2.4. The Cost-reward Analysis model

The Cost-reward Analysis model (Piliavin et al., 1969) is based on three assumptions: (a) potential benefactors are motivated to be self-benefitted (i.e., are egoistically motivated); (b) they are motivated to minimize the costs and maximize the rewards of their actions (economic perspective of motivation); (c) in their analysis of the costs and rewards, potential benefactors are relatively rational. The analysis process consists, basically, in pondering the probable circumstances, costs, and rewards that will result from the different behaviors that the potential benefactor has available (including of course not helping a person in need). Some examples of costs possibly involved in a prosocial behavior may be the time to spend helping, the interruption of other activities that the potential benefactor could be doing, and the danger implicated in helping; at the same time, costs for not helping may also be considered, like feelings of guilt, shame, or public censure. The rewards, on the other hand, may include the thanks from the beneficiary, money, avoiding guilt, or the pleasure or joy of having helped. It is assumed that the decision produced from this analysis is the best, most beneficial compromise between costs and rewards for the potential benefactor.

The model was developed considering the possible intervention of a potential benefactor in an emergency situation. It defines three conditions under which such assistance to a person in need is more likely: (a) the higher the aversive arousal, the higher the likelihood of helping (Dovidio, 1984; Gaertner & Dovidio, 1977); (b) the higher the sharing of feelings of proximity, similarity, or of a common group identity, the higher the likelihood of helping; and (c) the likelihood of helping is also higher when the implicated costs are considered lower than the rewards.

If the costs are too high – for example, if a potential benefactor must touch the blood of a person who is bleeding in order to help (Piliavin & Piliavin, 1972) – people may choose other means of reducing their own aversive state (e.g., by abandoning the place where the person in need is; e.g., Dovidio, et al., 1991). Usually, people put a higher weight on the costs than on the rewards.

Dovidio and Penner (2003) suggest that, in this model, having more positive attitudes and feelings for the person in need (like when this person is a relative or a friend) may increase the costs of not helping (e.g., more intense guilt), reduce the costs for helping (e.g., less anxiety because of knowing better how the potential beneficiary will react to the help), and increase the rewards of helping (e.g.,

23

attributing more importance to the gratitude expressed by the potential beneficiary) – making the prosocial behavior more likely to occur.

3.3. The theory of prosocial behavior with an altruistic motive – The Empathy-Altruism Hypothesis

The Empathy-Altruism Hypothesis is the only theory that postulates that the potential benefactor may act prosocially with an altruistic motivation – i.e., with an ultimate motive of relieving the potential beneficiary's need, or increasing his welfare – and is supported by plenty of evidence (Batson et al., 2002; Coke, et al., 1978; Dovidio, et al., 1990; see Batson, 2011 for a review). The reader should be reminded that relieving another person's need or increasing his welfare may be an ultimate motive whose attainment may produce self-benefits as unintended/unanticipated consequences (see The ultimate motive – providing pleasure or welfare to a target beneficiary in Chapter 3). This theory specifies that when perceiving a person in need, the subject experiences an emotion that is "*congruent* with the perceived welfare of another person*", which can be, for example, sadness or pity for someone who is upset or distressed, or empathy[1]. Empathic concern, a general designation for this emotional response, is considered "other oriented", meaning that it involves feelings *for* the other person – like sympathy, compassion, tenderness, or sorrow (Batson, et al., 2011). This emotion will elicit, in turn, the motivation to help the person in need.

Several experimental studies have shown that inducing empathy in a subject increases his likelihood of behaving prosocially toward another person. This is true even when it is easy for the potential benefactor to abandon the situation with the beneficiary, and thus avoid prolonging any aversive state that the former may be experiencing (Batson, et al., 2008). Indeed, there has been a contention over the years with some researchers claiming that empathically induced helping occurred only with egoistic motives (Cialdini et al., 1987; Schaller & Cialdini, 1988; Smith, et al., 1989), while others refuted such claim basing on methodological limitations found in those studies (Batson et al., 1988; Dovidio, et al., 1990; Schroeder, et al., 1988). Ultimately, one of the most representative researchers of the egoistic stance and his colleagues acknowledged that the existence of an altruistic motivation "does appear to have won the war in important respects" (Cialdini et al., 1997, p. 482).

Empathy may be activated or increased by valuing the welfare of the person in need (Batson, et al., 2007), by acknowledging kin, similarity, or proximity with that person (Cialdini, et al., 1997), or by sharing emotions, feelings, or sensations with him/her (de Vignemont & Singer, 2006).

Empathy or empathic concern may not be necessarily the only condition for altruistic motivation. Other factors have been studied as causes for behaving prosocially with an ultimate motive of increasing another's welfare, such as having an "altruistic personality" (Oliner & Oliner, 1988) and

---

[1] Although the term "empathy" has been used by scholars, over time, with different meanings (Batson, 2011).

acting under principled moral reasoning (Kohlberg, 1976). It is still not clear, however, whether these are sources of altruism, or rather other instrumental means for attaining egoistic ultimate motives, like to maintain one's positive self-concept or to avoid guilt (Batson, et al., 1986; Carlo, et al., 1991; Eisenberg et al., 1989).

People may avoid feeling empathic concern in order to avoid the resulting altruistic behavior. For example, Shaw, Batson, and Todd (1994) showed that participants aware they would be given an opportunity to help a target, chose to hear an empathy-inducing version of an appeal by a homeless man for help (vs. a non-empathy-inducing version) less often than those unaware of the upcoming opportunity.

Empathy may not always to produce prosocial effects. People may act against the collective good to benefit a person for whom empathy is felt (Batson, Ahmad, et al., 1999; Batson, et al., 1995), and even violate their own moral principles (Batson, et al., 1995).

3.4. The egoism/altruism debate

The existence of an altruistic motivation has been questioned and originated a long debate.

It is the case that a potential benefactor can be motivated to increase the potential beneficiary's welfare, and when the potential benefactor actually performs the behavior, or after completing it, he feels himself pleasure or some positive affect, and/or his own welfare increases, because of simply attaining a valued and desired state. Thus, an apparent altruistic motivation may have an egoistic result.

Given this, is it possible to behave with an ultimate motive of increasing the welfare of another person, or is any prosocial behavior always a means to obtain one's pleasure, satisfaction, or welfare? This question was named the altruism question, and has been extensively discussed (e.g., Batson & Powell, 2003).

Some scholars assume that, since being successful in attaining a motive always results in some positive affect for the subject, such effect *should* or *must be* assumed as an ultimate motive by him, before performing the behavior – even if unconsciously. The logic is that because the hedonic effect always occurs, it should always also be anticipated by the actor, thus qualifying any prosocial act as egoistic. People help because they know, even unconsciously, they will benefit in the end. For these theorists the egoistic motivation is universal (Batson & Powell, 2003).

Others, however, disagree with this assertion, arguing that whether or not the attainment of some self-benefit indeed always occurs, it does not mean that it was the ultimate motive for having the behavior, or rather could be only an *unintended consequence* of having accomplished the motive. Thus, the benefactor's positive affect arising from acting prosocially can be, in this perspective, an unintended consequence of the motive of promoting the satisfaction of another's need and increasing

25

their welfare. This is the perspective that advocates the possibility of an altruistic motivation (Batson, 2011; Batson & Powell, 2003). From this we can deduce that even if, theoretically, the potential benefactor knew in advance he was incapable of experiencing any kind of positive affect resulting from behaving prosocially, he would still perform the behavior. Research suggests that, when people are empathically aroused, this actually happens: participants of studies help even when they are led to believe that helping cannot improve their mood (e.g., Schroeder, et al., 1988).

Chapter 4. Withholding prosocial behavior

Now that we know why people may help, we are more able to understand why people may withhold help. The way people use information about others when deciding to help is rarely investigated. Why, and how does behavioral information influence helping decision making? Why people consider that some deserve more help than others? The main goal of the present thesis is to analyze how the decision of whether to help a person can be influenced by aggressive behaviors (specifically, hitting, hindering, and threatening behaviors) by that person toward a third person in the past.

The literature provides two sets of answers for these questions: one set that does not consider the role of a potential beneficiary's behavioral information in the potential benefactor's decision – which although not responding directly to my questions, helps to understand help withholding –, and another set of answers that considers such information. Let us analyze each set.

When a potential beneficiary's behavioral information is unknown or irrelevant

One reason for withholding help can be *not perceiving or knowing the need of help* (Latané & Darley, 1970). This situation is not explored in this work, since it does not trigger any decision process involving help. However, because of being an essential step to activate prosocial behavior, in the empirical studies that follow in the next section we measured participants' perception of a target's need.

Some people may feel *uncertainty* regarding how to help, or feelings of *inability* to complete the needed behavior of helping. Possibly due to a difference in such feelings, students were less likely than nurses to help a workman who had fallen off a ladder (both students and nurses having reported in a posterior questionnaire that they felt strongly that they should help; Cramer, et al., 1988).

Help may also be withheld if the potential benefactor believes that helping can bring *unintended consequences*. These anticipated, undesired states may compete with motives for helping, and end up deactivating them, and cancelling a possible prosocial behavior (see Motivation and willingness in Chapter 2.2.). Because the potential benefactor may thus be motivated to avoid anticipated states, if he experiences an aversive arousal and those states are pondered rationally as costs of helping, he may then be performing a cost-reward analysis that results in help withheld (see 3.2.4. The Cost-reward Analysis model; Piliavin, et al., 1969). Jones (1970) showed that undergraduates who were told that a peer in need would continue requesting help in the future, from them, were more likely to refuse helping than those who were not told. Knowing that they would have to help again in the future may have dissuaded some of those who refused.

Behaving prosocially due to an affective response, when facing another person in need, is the fundamental process on which the aversive arousal theory, the negative state theory, the cost-reward analysis model, and the empathy-altruism hypothesis are based. In any of them, it seems clear that if

the potential benefactor does not experience such an affective or emotional arousal, he will not help (see Affective or emotional arousal and motivation in 3.2.3. Avoidance of aversive situations or punishments; e.g., Piliavin et al., 1982). But would it be possible that the potential benefactor does not experience a *sufficient level of affective response* to trigger the mechanisms proposed by those theories that lead to help? If there is a threshold of affective activation, it is reasonable that people might withhold help because of not reaching that threshold. This may be considered a form of apathy. According to the cost-reward analysis model, for example, the higher the aversive arousal, the higher the likelihood of helping (Dovidio, 1984; Gaertner & Dovidio, 1977).

When a potential beneficiary's behavioral information is known or relevant, and negative

We saw some evidence regarding people withholding help when not possessing any behavioral information about the potential beneficiary (or when such information is irrelevant). But let us now turn to when people know something about the potential beneficiary that can be relevant for deciding to help. From what is known of the different theories explaining prosocial behavior, we can postulate some ways in which behavioral information may influence helping decisions.

Undermining trust in the potential beneficiary's need – it is possible that a potential beneficiary' bad reputation casts doubt on the potential benefactor's belief in the former's need. If a potential beneficiary is thought of as manipulative, when he drops his books in front of the potential benefactor, the potential benefactor may interpret such event as contrived to obtain help for some reason other than related to the books, affecting his motivation to help. People tend to help less frequently individuals with a bad reputation (Wedekind & Milinski, 2000; see Social and behavioral resources – reciprocity behaviors in Chapter 3.2.2.).

Undermining beliefs in the likelihood of self-benefit – egoistically motivated people may help someone in need because of believing that their behaviors will result, somewhere in time, in self-benefit. Thus, like with any other motive, people *act because they believe* that some behavior will produce a certain state that is valued (see 2.2. Motive/goal, and motivation). However, acquiring some particular behavioral information about the potential beneficiary may invalidate that belief; some information may make people cease to believe (or prevent them from ever believing) that helping will produce a certain state/a self-benefit. For example, if the potential benefactor who would help with the expectation of reciprocity receives the information that the potential beneficiary has previously not reciprocated help to a third person, he would possibly question whether the potential beneficiary would not reciprocate to him as well. Additionally, helping someone about whom one has learned negative information might be considered less likely to gain a benefactor social approval, or even lead to social censure. For example, in public good games, non-cooperative behaviors are disapproved and generate negative emotions by cooperative individuals (Fehr and Gachter, 2000, 2002; Chapter 1). As

28

mentioned, behavioral information may also undermine beliefs about possible anticipated consequences of withholding help. The potential benefactor may help believing that the alternative behavior, withholding help, would bring an aversive situation – for example shame, or disapproval of others (see 3.2.1. Characterization of self-benefits). But would others disapprove of withholding help to a person who usually has an antisocial demeanor? In this case the potential benefactor may no longer believe that withholding help will cause others' disapproval. Again, an opposite belief may even be activated: that if he helps, he will be disapproved by others.

Devaluing self-benefits involving the potential beneficiary – also only happening with egoistically motivated people, this phenomenon can be seen as a particular case of withholding help because of not valuing consequent states of helping – one in which such states involve the potential beneficiary. Let us return to our antisocial potential beneficiaries. It does not seem likely that a potential benefactor would help a drug addict in order to obtain his esteem or approval. As it seems questionable that those potential benefactors who help expecting to be reciprocated later would want to be benefitted by potential beneficiaries with the referred past behaviors. Indeed, often times, when knowing about potential beneficiaries' condemnable behaviors, people do not even want contact with them (e.g., Ahern, et al., 2007) – let alone help them with a goal of reciprocation or thanks. It may also not be attractive to obtain a reputation of having helped an antisocial beneficiary, in order to later benefit from third parties' prosocial behaviors (in fact, regardless of reason; see Social and behavioral resources – reciprocity behaviors in Chapter 3.2.2.).

Conditions to reciprocate – people who were helped may not, in some conditions, feel compelled to reciprocate the received help (despite the universality of the reciprocity norm). In his seminal article, sociologist Alvin Gouldner (1960) identifies the conditions according to which the compliance with the norm and reciprocation may vary. One of those conditions is the intensity of the beneficiary's need when he was helped in the first place. If the intensity is considered by the beneficiary as low, it is possible that he does not consider the help he was given needed – i.e., he may consider that the benefactor should not have helped, because he did not need. To reciprocate would then mean to validate, or to acknowledge this wrongly given help, which the beneficiary does not want. Thus, unnecessary given help is not compatible with the norm and so no debt is contracted. Another determinant of reciprocation is the perception the beneficiary has of the amount of help he received from the benefactor, relative to his total resources. Gouldner (1960) suggests that willingness to reciprocate may be high if the benefactor could ill afford the given help; but what if the given help represented an insignificant effort or resource of the benefactor? It seems reasonable to think that it can undermine the beneficiary's willingness to reciprocate (and to comply with the reciprocity norm). A third relevant determinant to reciprocate pointed out by Gouldner (1960) is the motivation of the benefactor to help, namely whether he had the reciprocating behavior or other self-benefit in mind

29

when helping. To perceive or suspect a benefactor's egoistic motive may also compromise the compliance with the norm, preventing the beneficiary from reciprocating.

Particular effects in a cost-reward analysis – it is possible that a potential beneficiary's behavioral information suggests additional costs to potential benefactors who perform cost-reward analyses before helping (see 3.2.4. The Cost-reward analysis model; Piliavin et al., 1969). They may think for example that if they help a moral offender, like someone who threatened or manipulated others, they will later experience the unintended consequence of feeling guilt or regret. Or, if such help is public, that some close people will censure them. The weight placed on some costs may also be increased: contemplated costs like time and energy may be aggravated when considering a potential beneficiary with a reprehensible past. For example, a teacher may be less willing to help an undisciplined student doing his homework than a dedicated one, expecting more value made from the help given. On the other hand, the behavioral information may also affect the rewards: predicted self-benefits may not be believed to occur or lose their value (see previous topics).

Preventing empathy – as explained before, not feeling enough empathy may inhibit helping, regardless of any previous behavioral record of the potential beneficiary. It can possibly be due to individual characteristics, personality, or circumstantial factors. But empathy may also be affected by some behavioral information of the potential beneficiary, thereby preventing helping behavior. For example, perceived dissimilarity in personality, values and interests with the potential beneficiary may trigger less empathy than perceiving similarity with him (Krebs, 1975).

Punishment and negative behavioral information

Negative behavioral information may influence helping decisions also by triggering punishment behaviors – a person's negative behavioral information has been shown to cause punishment through removing resources, administered by others. Indeed, Fehr and Gachter (2000; 2002) demonstrated that, in public good games, information about players' previous non-cooperative behaviors elicited negative emotions in other players which in turn made them punish the former, by removing their resources. Players may adopt this punishment to reinforce prosocial norms, to increase cooperation, and/or to increase their own or other peoples' payoffs/resources (see Chapter 1. Introduction; see Figure 4).

Figure 4

Cognitive Process in which Behavioral Information causes Punishment Behavior
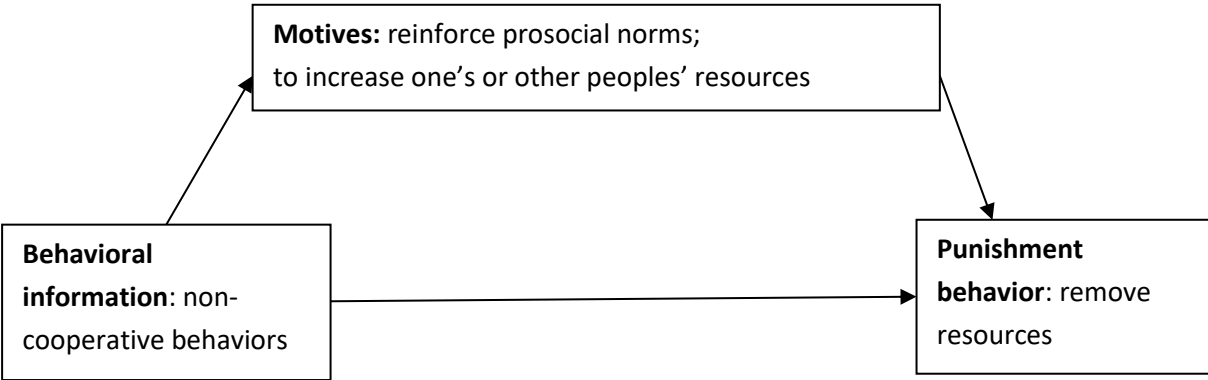
Other motives for punishment may be inferred from criminal justice theories, since people resort to these to judge criminal cases.

Just deserts punishment theory

In general, retributionist punishment seeks an equivalence between a committed offense and the punishment, in a way that the offender loses what the victim has lost (Hudson, 2003). For most of the twentieth century, until about 1970, retributivism was considered by criminologists as simple vengeance, as Barbara Hudson puts, "an unprogressive basis for punishment", serving little to the criminal justice system; only over the last quarter of the twentieth century into the early twenty-first a reemergence has occurred, with great strength and influence, under the name of "just deserts" (Haist, 2009; Hudson, 2003). There was a shift in the rhetoric of scholars from victims or society taking severe vengeance on the criminal to getting what he or she *deserved* for having committed the crime, i.e., "a punishment in accordance with the severity of the crime and his or her moral culpability." (Haist, 2009).

Herbert Morris, one of the pioneers of this reformed retributivism, reasoned that the law created equal rights and responsibilities for all members of the society; and when an individual committed a crime, he or she violated this equal distribution by continuing to benefit from others not committing crimes and being responsible, but violating his own responsibility of not committing crimes (Morris, 1981): "a person who violates the rules has something others have – the benefits of the system – but by renouncing what others have assumed, the burdens of self-restraint, he has acquired an unfair advantage" (p. 95). With the committed crime, the offender obtained an unfair advantage which the law-abiding deny themselves by obeying the laws (Hudson, 2003). In this sense, a just deserts

punishment would be society removing the advantages the offender had seized unfairly, and restoring the proper balance of benefits and burdens of the law (Finnis, 1972; Murphy, 1979; Sadurski, 1985).

This view was, however, further criticized by scholars for the simple reason that there are crimes which non-offenders would never be tempted to commit (e.g., rape, murder), and thus would never benefit of any unfair advantage possibly resulting from those. The just deserts justification or function of removing an unfair advantage would then be only applicable to some categories of offenses (Duff, 1986).

Other authors do not agree on a more specific definition. Some say that it is considered a valuable end in itself, needing no justification, regardless of its consequences for either the offender or the society (Carlsmith, 2006; Carlsmith, et al., 2002; Darley, et al., 2000). It is also considered an expression of retaliation ("an eye for an eye") or vindictiveness (Gerber & Jackson, 2013). It does not aim to change future behavior, and any consequences of punishment are irrelevant for its assignment (Carlsmith & Darley, 2008; Darley & Pittman, 2003); the punishment is still needed and effective if the offender expresses regret for the offense, and/or would continue committing offenses (Hudson, 2003). In this respect, retributionist theory assumes that although a criminal chose to do a wrong action, he/she can still choose a right action in the future; breaking a particular rule does not mean that the he/she rejects the law altogether (Hudson, 2003). Other authors understand crimes as violations of society's rules that threaten moral values and norms (Carlsmith, 2006; Miller & Vidmar, 1981; Vidmar & Miller, 1980), and find in a just deserts punishment the symbolic justifications of: (a) reasserting the community commitment to those threatened norms and worldviews, and reasserting the obligation to obey social rules (Darley & Pittman, 2003; Hamilton & Sanders, 1988; Miller & Vidmar, 1981; Vidmar & Miller, 1980); Darley and Pittman (2003) argue that societal norms and worldviews are typically learned in a person's early development, and, as probable fundamental aspects of comprehending the reality, are likely to be defended with vigor when threatened; (b) restoring social consensus about the rules and values that have been violated, not only with other people (e.g., supporters of the punishment), but also with the offender him/herself, who may realize the offense and feel remorse (Vidmar, 2000; Vidmar & Miller, 1980). By suffering the punishment, the offender shares, or is induced to share the view that his/her offense was wrong, thus being allowed to rejoin the society's moral consensus (Hudson, 2003); the purpose of punishment would then be to reconnect the offender with society's values (Nozick, 1981); (c) restoring social balance (e.g., Carlsmith, et al., 2002; Weiner, et al., 1997); and (d) communicating moral values to society (Gerber & Jackson, 2013), and publicly expressing disapproval of a perpetrated behavior; the degree of severity of the punishment applied indicates the degree of that disapproval (Hart, 1968).

Regarding the function of restoring the social balance, Wenzel and Thielmann (2006) explain, referring themselves to a just deserts punishment:

By breaking the rules, the offender has taken advantage of the victim and the society, has

expressed contempt for the victim and society, and has assumed a position of superiority and

power over victim and society (…). The punishment is meant to reverse this situation. (p.451)

In any case, to achieve any of these goals, the "perpetrator should suffer in proportion to his or her wrongdoing" – the principle of moral proportionality (e.g., Banks, 2017; Carlsmith, 2006, p. 437). Thus, punishment should be commensurate (i.e., proportionate to the seriousness or gravity of the offense): the most serious punishment should be given to the most serious offense, and for the remaining offenses punishment severities should be graduated in a scale according to the offense seriousness (a logic that is also known as tariff sentencing: for each crime there is a specific price to pay; Hudson, 2003). Notwithstanding, studies conducted both with judges (Hogarth, 1971) and laymen (Warr & Stafford, 1984) show that retribution is associated with harsh punishment. This suffering can be accomplished by reducing the offender's status and power relatively to the victim and the society (Vidmar, 2000; Vidmar & Miller, 1980), which restores the social balance (Wenzel & Thielmann, 2006).

According to another conceptualization, there may be two types of retribution, depending on its more specific goal (Von Hirsch, 1976): a) retribution as just deserts, in which the offender must pay back for the offense committed, proportionally, and thus positive and negative experiences are distributed and social balance is restored (Weiner et al., 1997); and b) retribution as revenge, aiming not only to restore social balance, but also retaliation, i.e., the offender's suffering (as well as the pleasure of seeing such suffering; Banks, 2008; Ho et al. 2002; Weiner et al., 1997). While in the former type it is the offender who compensates for the harm, in the latter that compensation is accomplished by the society (Finckenhauer, 1988). Having conducted an online survey, Gerber and Jackson (2013) found empirical support for this distinction: a two-factor model of retribution showed a better fit than a one-factor model. Specifically, the dimension of retribution as revenge involved the subdimensions of getting even and making the offender suffer and was precisely defined as "a desire to get even with the offender by making him/her suffer", whereas the second dimension comprised compensation and proportionality, having been defined as a "desire to restore justice by allowing the perpetrator to compensate proportionally for the harm done". Darley and Pittman (2003) use a simple example to illustrate this distinction:

If a thief, eventually caught, were to compensate the victim immediately for the full monetary

costs of the theft, people would not consider that this was a sufficient infliction on the

perpetrator to close the matter. Instead people feel that suitable punishment must be inflicted

on the perpetrator by the judicial system, a punishment in proportion to the moral gravity of

the offense committed. This is not accomplished simply by requiring the perpetrator to compensate the victim for damages. (p.326)

In a review of literature, Gerber and Jackson (2013) noted that retribution is measured using items that capture different dimensions such as deservingness, morality, retaliation of the offense, proportionality, social cohesion, and concerns about justice.

Punishment should primarily attend the moral gravity of the offense, and secondarily the characteristics of the offender (Miethe & Lu, 2005). The likelihood of recidivism "should not influence the sentence passed, or the amount of punishment actually undergone" (Hudson, 2003). In court, sentence reductions are allowed, however, for first offenders; this is because besides considering the seriousness of the offense, the just deserts sentencing also considers the culpability of the offender, i.e., his or her intentionality in committing the harm. This means for instance that "an accidental or coerced act would not carry the penalty of a freely chosen, fully willed action"; a first offender may not realize the consequences of his or her actions, nor the impact on victims (which is less likely in the case of recidivists; Hudson, 2003).

Examples of pure just deserts punishments include requiring a defrauder to pay the exact stolen amount, or executing an assassin.

Utilitarian punishment theories: incapacitation and deterrence

This perspective is distinct from utilitarian or instrumental theories, which encompass the incapacitation and the deterrence stances (Carlsmith, et al., 2002; Darley, et al., 2000). According to utilitarian theories, punishment should only be used to prevent or reduce possible future harms, by minimizing the likelihood of future rule breaking (Darley, et al., 2000). Utilitarian theories (sometimes referred as consequentialist) derive from the utilitarian movement in philosophy, in which all behaviors should be aimed at improving happiness in the world (Carlsmith et al., 2008). Despite resulting in a decrease of happiness of the targeted person, a utilitarian punishment is used and expected to accomplish the more distal goal of increasing the society's happiness (Carlsmith et al., 2008). In effect, utilitarian theories are only concerned with how punishment may affect future actions and society's future happiness; thus, "punishment should only be administered if it results in an overall benefit to society" (Haist, 2009).

In the case of the incapacitation theory, such harm avoidance is accomplished by removing the means of offending (e.g., disbarment), or by preventing any contact between the offender and potential victims (e.g., through incarceration or deportation) – i.e., the opportunity to cause harm is restricted or eliminated, at least for a relatively long time (Carlsmith et al., 2008; Miethe & Lu, 2005; Nagin, 1998). Incapacitation is also designed to decrease the physical capacity of an individual to

commit potential offenses (Miethe & Lu, 2005). One favorable factor for the use of this punishment is the belief that a significant number of crimes are committed by a relatively small number of repeated offenders (Greenwood & Abrahamse, 1982). A central task of the incapacitationist system is then to identify those offenders who are likely to recidivate (Robinson & Darley, 1997).

Over time, it has for instance resorted to banishment to the wilderness by primitive tribes, to exile in the ancient Greece and by the English to colonies in the seventeenth and eighteenth centuries, and to confinement to forced labor contexts (e.g., forced servitude in military campaigns) and correctional centers; nowadays, incapacitation instruments include the use of electronic ankle bracelets (that, for example, automatically inform authorities if the offender trespassed an interdicted area), child harnesses that limit the movement of children in public areas, and chastity belts or castration to impede sexual activity (Carlsmith et al., 2008; Miethe & Lu, 2005).

Beyond varying in their nature, incapacitation sentences may also vary in their limits or scope. For example, after serving a sentence in prison, an offender may be temporarily placed in a halfway house, in a program designed to ease the adjustment to conventional life (Miethe & Lu, 2005).

According to Darley, Carlsmith and Robinson (2000), the incapacitation stance suggests that the primary determinant of a punishment should be the perpetrator's likelihood of committing future offenses (recidivating).

Incapacitation sentencing may bring some ethical and practical problems, such as dealing with both "false positives" (i.e., punishing people who are not high-risk offenders) and "false negatives" (i.e., releasing high-risk offenders; Miethe & Lu, 2005). Even though certain offenders commit crimes at very high rates, predictions about these people's further offenses may not always be accurate, inevitably resulting in a number of incapacitated persons who would not actually reoffend should they be left free; moreover, it is argued that, in principle, offenders should be punished for what they have done, not for what they may do, or not, in the future (Banks, 2017). Peter Schmidt and Ann Witte created a model to predict recidivism (1988) and found that their best model was too inaccurate to be used in assigning people to prison or freedom (e.g., their model falsely identified 47% of offenders as going to recidivate, i.e., people who did not recidivate in fact). This led scholars to advocate that offenders should be incapacitated only when there is reliable information indicating a high probability of reoffending (Morris, 1994). Another obstacle to incapacitation sentencing is that, because crimes are commonly committed by groups of offenders, incarcerating one member does not impede the other members from continuing on commit crimes (or recruiting a new member; e.g., drug dealers are easily replaced), compromising the goal of the punishment; in this sense, incarceration may only affect the identity of the offender, not the volume of offenses (Nagin, 1998).

Deterrence theory assumes that citizens are rational actors, who choose courses of action to maximize pleasure and minimize pain; it works by inducing a change in the perceived costs and benefits

35

of a possible harm so that it becomes an unattractive option, for past or potential offenders (Carlsmith & Darley, 2008; Miethe & Lu, 2005; Nagin, 1998). Specifically, the punisher believes that, by punishing a perpetrator, other potential offenders who know about the punishment will associate (and anticipate) the additional cost (and consequence) of being punished (and thus of feeling pain, dissatisfaction, or other causes of negative affect) with committing an offense. Thus, when anticipating behaving offensively, the predicted consequences will inhibit the act (or reduce its likelihood; as long as the punishment is higher than the offense's payoff). The punisher expects the potential offender (a rational actor) to run a cost-benefit analysis before behaving offensively. Forms of punishment such as fines are designed to induce this reasoning (Carlsmith, et al., 2002).

In this sense, punishment should be public or publicized[2]; a private deterrence-motivated punishment would be irrational or immoral, since it would create the unhappiness of the offender without any deterrent effect in other potential offenders. Although deterrence punishment is traditionally aimed to prevent other potential offenders from committing harms, not the person who is punished (which is called "general deterrence"), such a motive is also activated for preventing a past offender from repeating harms in the future ("individual or specific deterrence"; Carlsmith, et al., 2008; Hudson, 2003; Miethe & Lu, 2005).

The efficacy of a deterrence punishment may be assessed through indicators, such as the frequency of re-offenses (an increasing rate of offenses indicates that the punishment was not sufficiently deterrent), and the offense detection rate (e.g., if most committed offenses are not detected then punishments aimed to prevent them naturally were not effective; Carlsmith, 2006). Deterrence punishment is also assumed to be more effective when applied to offenders who have a "low level of commitment to deviance as a livelihood" (e.g., people who are not professional criminals; Miethe & Lu, 2005).

There are certain similarities of the deterrence punishment rationale and operant learning (Cavender, 1979). Both theories rely on the assumption that people are motivated, in their behaviors, to derive pleasure and to avoid pain. A criminal sanction used with a deterrence goal is equivalent, in the operant learning theory, to a punishment that follows a behavior one wants to discontinue or extinguish. Other factors that influence the efficacy of a deterrence – or operant learning – punishment are: a) the estimated likelihood (by the offender or other potential offenders) of occurrence of the punishment caused by a certain behavior – called "certainty" in the deterrence theory, and "schedule" in the operant learning theory; b) the swiftness with which the punishment follows the offense – which

---

[2] (from a strictly logical and extreme point of view, according to deterrence, even true guilt or responsibility for the offense is irrelevant, as long as other potential offenders recognize that responsibility, and associate the punishment to a possible offense they may consider doing; thus, according to this, it would be legitimate to punish an innocent, since it could bring a societal benefit, i.e., it would have utility; see Carlsmith et al., 2008).

is the "celerity" in the deterrence theory, and "immediacy" in the operant learning; and c) the degree of pain or suffering caused by the punishment – the "severity" in the deterrence theory, and "intensity" in the operant learning theory (Azrin & Holz, 1966; Bentham, 1973; Cavender, 1979).

Fines are mostly used with a deterrence rationale. Other methods include using breathalyzers (device that measures the concentration of alcohol in a person's blood), injunction or restraining orders, revocation of licenses (e.g., driver's license), and economic sanctions such as boycotts.

Note that a potential offender might be dissuaded from committing an offense because of having in mind the possible outcome of being imprisoned, i.e., incapacitated. But it also happens the other way around: certain punitive measures with a primary deterrent intent may have also incapacitative collateral effects: ostracism, adverse publicity and censorship may restrict opportunities (Miethe & Lu, 2005).

Just deserts versus utilitarian punishment

Thus, one can conclude that the goal of just deserts punishment is to repair a harm done, focusing on the past, whereas the goal of an incapacitation or a deterrence punishment is to protect people and prevent harm in the future.

Frequently, criminal punishment fulfills more than one goal effectively (Carlsmith et al., 2008). For instance, imprisonment is considered to attain both the incapacitation and a just deserts goals (Carlsmith, et al., 2002). In the same sense, assigning a just deserts punishment may have a deterrent effect (specific and/or general), but only as a byproduct of a just deserts orientation rather than the goal of the punishment (Darley & Pittman, 2003). There are other utilitarian perspectives that we do not use in our research[3].

Empirical findings also revealed that people, when asked to assign punishment, more often seek out information relevant to a just deserts than to a utilitarian sentencing (Carlsmith, 2006). People also seem to punish more often with a just deserts motive than with utilitarian motives (Carlsmith et al., 2002; Darley, et al., 2000), at least in the US; in Japan, for example, people seem more favorably disposed to deterrence and other utilitarian motives (Hamilton & Sanders, 1988). But these results may differ on the justice orientation people adopt. Brickman et al. (1981) suggest that people endorse different justice criteria depending on the orientation they activate, which can be either: a) a microjustice orientation, centered in the attributes of the individuals, without considering the society's

---

[3] These include the rehabilitationist justification, which consists in subjecting wrongdoers "to a set of training and attitudinal adjustment procedures that will reeducate them so that they have the resources and motivations to become productive members of society" (Darley & Pittman, 2003). However, researchers consider this justification to be only tenuously connected to the ordinary concept of punishment. Other utilitarian justifications such as restoration and restitution, were not included in our research, as they are less associated with citizens' intuitions of justice (Carlsmith & Darley, 2008).

needs; or b) a macrojustice orientation, focused on the well-being of the society, and in which the deservingness of a particular outcome for an individual is secondary. Following this conceptualization, Miller and Vidmar (1981) found that people who are macrojustice oriented triggered a utilitarian mindset, whereas those who are microjustice oriented triggered a retributive mindset. This view is consistent with concerns with the society as a whole being represented by punishing with a general goal of prevention of further offenses, and concerns with the offender (and victims) being represented by the goal of restitution (Oswald et al., 2002). Indeed, scholars differentiate sentencing goals depending on the target of the justice considerations – the society in general or the offender (e.g., Miller & Vidmar, 1981), and, Oswald et al. (2002) argue,

> Only rarely can one assume that a particular punishment is accepted by all those concerned, and thus the choice of a particular punishment means, in most cases, also a decision as to which target's concern is to be given the most weight. Thus, a decision in favor of a lengthy therapy may serve the needs of the offender well but not those of society, given the high costs (p.88).

Based on the Temporal Construal Theory of Liberman and Trope (1998), Carlsmith (2008) suggests that temporally proximate decisions (like a specific sentencing task) elicit a microjustice orientation in individuals, and their retributive punishment, whereas distant decisions (like when an individual is asked to recommend general goals for punishment or sentencing policies) activate a macrojustice orientation and the application of utilitarian punishments.

Evidence seems to suggest that punishing is more intuitive – automatic, rapid, and possible of co-occurring with other processes – rather than a completely reasoned decision (Carlsmith, 2008). Research shows that people quickly judge if certain behaviors are morally wrong, and/or punishment deserving, but are unable to describe the reasoning that led them to such conclusion (e.g., Haidt, 2001; Hauser et al., 2007).

Measurement of punishment motives

Empirical research on punishment has shown that the use of scales to measure people's agreement with these punishment motives was inaccurate because all motives sounded socially desirable, and people tended to agree with them all (Anderson & MacCoun, 1999). It is also known that people are sometimes unaware of the actual reasons that drive their decisions (Nisbett & Wilson, 1977). Because of this, Carlsmith and colleagues (e.g., Darley, et al., 2000) have used a research paradigm in which participants are presented with descriptions of harm-doing containing variable information that should be relevant for punishing with specific motives (i.e., some descriptions included information relevant

38

for punishing with a just deserts motive, and not relevant for punishing with an incapacitation motive – and vice-versa). Participants are then asked to assign a punishment severity for the harm-doing, and the determinant motive that leads them to punish is indicated by the severity of punishment assigned given a specific description. Literature shows that the seriousness/magnitude of the harm committed is an important factor for assigning a just deserts motive, and that it is less relevant for the desire to incapacitate; on the other hand, the likelihood of recidivating by the offender (as well as his history of offenses) seemed important for activating an incapacitation motive in studies, but not for activating a just deserts motive (Darley, et al., 2000). This paradigm was adopted in the studies we conducted, using harm magnitude to manipulate just deserts motive, and a target's past offense history to manipulate incapacitation and deterrence motives in our studies.

Behavioral information may thus cause punishment behaviors, which can assume different forms (e.g., removing resources or deterring) with different motives (e.g., increasing one's resources, or protecting oneself).

Chapter 5. Current research

The present thesis proposes that withholding help may be a punishment behavior triggered by particular behavioral information about a potential beneficiary. Such inhibition, as a punishment, may be motivated by any of the aforementioned motives for administering punishments (see Figure 5). In other words, in order to attain punishment goals, help cannot be provided (e.g., if I want to retaliate, i.e., inflict proportional suffering, I cannot help – help would prevent the person from suffering; see 2.2. Motive/goal, and motivation).

Figure 5

Cognitive Process in which Behavioral Information may cause Particular Punishment Behaviors with Different Motives

Punishment motives
(of a potential benefactor) – Examples:
- Reinforce prosocial norms to increase one's resources
- Modify transgressive behaviors
- Retaliate
- Reassert commitment to norms, worldviews, and to obey social rules
- Restore social consensus about rules and values that were violated
- Communicate moral values to society
- Prevent or reduce future harms or rule breaking
- Prevent contact between the offender and potential victims
- Remove means of offending
- Deter by changing the costs and benefits of possible harm

Negative behavioral information
(of a potential beneficiary) – Example: non-cooperative behaviors

Decisions involving behaviors / punishment behaviors
(of a potential benefactor) – Examples:
- Removing resources
- Inflict proportional suffering
- Incarcerate; deport
- Suspend professional activity
- **Withholding help**

Negative behavioral information about the potential beneficiary may, then, trigger a punishment motive and consequent help withholding – in a potential benefactor who would otherwise help (whether with an egoistic or altruistic motive). That is, information about negative past behaviors of

40

the potential beneficiary may activate a punishment motive in the potential benefactor, inhibiting the activation of a motive to help – or competing with the motive to help (see 2.2. Motive/goal, and motivation).

As opposed to other proposed reasons for withholding help due to the recipient's behavioral information, withholding help as a punishment may be, theoretically, independent from any particular motive that would drive people to help. That is, it may occur whether the potential benefactor would help to attain some personal norm, expect some reciprocity behavior, increase self-esteem, reduce an aversive arousal, avoid sanctions, because of feeling empathy, with any egoistic, or altruistic motivation. Whatever theory would explain prosocial behavior in any situation, behavioral information triggering punishment motives causing help withholding may always happen.

In order to test this possibility, empirical studies with experimental methods were designed and conducted, as follow in the next section. In the Empirical Section we will describe one pretest and three subsequent studies.

EMPIRICAL SECTION – EXPERIMENTAL STUDIES

In the present thesis we present research examining whether participants would be less likely to help a needy target who behaved aggressively toward a third person in the past – testing Vaish, Carpenter, and Tomasello's (2010) findings with a young-adult sample – and extending it by manipulating punishment motives.

We conducted one pretest and three experimental studies to test our hypotheses. With the pretest we wanted to test our materials, procedures, and measures, which worked according to our goals. We then ran the main study, Study 1, whose unpredicted results and interpretations led to a follow-up study, Study 2. Here we made changes in the materials, additions in the measures, and tested new hypotheses. Results of Study 2 provided support for some of our hypotheses. In Study 3 we tested new hypotheses for explaining help withhold, which also corroborated some of our predictions and allowed a clearer picture of selectivity mechanisms of help.

In the following pretest we tested the materials, procedures, and measures with a small sample in order to make any possible adjustments that could benefit the design of the main studies.

5.1. PRETEST

Aims and Hypotheses

Based on the assumption that participants exposed to aggression information will want to punish the target by withholding help toward a person in need, and on Darley and colleagues' paradigm (2000; see Chapter 4), according to which assigning more severe punishments on a punishment severity item would confirm the activation of each motive (i.e., just deserts or incapacitation), we formulated custom hypotheses within the conditions, and tested them using planned contrasts. They are the following:

Participants would withhold help as a form of punishment with a just deserts motive (H1), if we obtained:

o A main effect of harm magnitude on the punishment severity, with participants in the high harm conditions punishing the target more severely than in the low conditions, indicating the activation of a just-deserts motive for punishing in the punishment severity item (H1.1.);

o A main effect of harm magnitude on helping, with participants in the high harm conditions helping less than those in low harm conditions (H1.2.);

o Participants in the high harm/low history condition helping less than those in a no-aggression condition (H1.3.).

Participants would withhold help as a form of punishment with an incapacitation motive (H2), if likewise, we obtained:

o   A main effect of the target's past offense history on the punishment severity, with participants in the high history conditions punishing the target more severely than in the low conditions, indicating the activation of an incapacitation motive for punishing in the punishment severity item (H2.1.);

o   A main effect of the target's past offense history on helping, with participants in the high history conditions helping less than those in the low history conditions (H2.2.);

o   Participants in the low harm/high history condition were also expected to help less than those in a no-aggression condition (H2.3.).

We did not include the punishment question in the no-aggression condition (as it did not make sense), thus no comparison on this item could be made.

The harm magnitude of the aggression (high or low) should not have a significant influence for punishing with an incapacitation motive because empirical findings showed that harm is less relevant for the activation of that motive. Conversely, the target's history (high or low) should not significantly affect a just deserts punishment because empirical findings have found history to be less relevant for that motive (Darley, et al., 2000).

Assigning a more severe punishment and helping less in a condition with a high harm magnitude of the aggression and in which the target committed no past aggressions (i.e. low in history) – hereafter high harm/low history – than in low harm magnitude and no-aggression conditions was unlikely to be motivated by an incapacitation motive, because that condition did not contain relevant information for activating such motive (harm magnitude has been found to be less relevant for punishing with an incapacitation motive, see Darley, et al., 2000). Rather, the high harm/low history condition included relevant information to punish with a just deserts motive (should participants wanted to punish the target with that motive). Thus, according to research (Darley, et al., 2000), participants in the high harm/low history condition were provided with information that was relevant for punishing with a just deserts motive (high harm magnitude), if they wanted, and information less relevant for punishing with an incapacitation motive (no past aggressions).

The same rationale, but inverted, applied to the condition in which the target had an history of aggressions (i.e., high history) that were low in harm magnitude – hereafter low harm/high history condition: this condition contained relevant information to punish with an incapacitation motive (target's past aggressions), and less relevant information to punish with a just deserts motive (low harm magnitude of those aggressions).

We also considered possible that some participants would want to punish the target with *both* a just deserts and an incapacitation motive, as follows:

43

Participants would withhold help as a form of punishment with *both* a just deserts and an incapacitation motive (H3), if we obtained:

- o Main effects of both harm magnitude and past offense history on the punishment severity variable (H3.1.);
- o Main effects of both harm magnitude and past offense history on the helping variable (H3.2.);
- o Participants in a high harm magnitude/high history condition helping less than those in the high harm magnitude/low history, low harm magnitude/high history, and no-aggression conditions (H3.3.).

Method

Participants

Forty-nine undergraduate students (aged 18-66; *M* = 20.12, *SD* = 6.95) participated in the study in exchange for course credit.

Measures

We adapted a paradigm previously used by Batson and colleagues (e.g., Coke, et al., 1978) in which participants are presented with a bogus pilot radio program resulting from a partnership between the university and a well-known radio station. This paradigm has consistently triggered helping behavior in listeners (Batson, 2011).

Participants were randomly assigned to one of five conditions: one control condition with no-aggression information and four conditions with aggression information in a 2 (harm magnitude: high vs. low) X 2 (history: high vs. low) between-subjects design. As in previous similar helping studies (e.g., Toi & Batson, 1982), only female subjects participated, because it was shown that empathy for the female target would be higher if participants were the same gender as the target.

Participants attended sessions individually. On arrival, they read and signed a consent form, and were told that the study would measure physiological responses. Two electrodes for skin conductance levels (SCL) were attached unilaterally to the proximal phalanges of the middle and third fingers of their non-dominant hands. This physiological measure was not central to the study and was not analyzed, having only been used as an additional means to conceal the study's purposes. Participants read the introduction to the study, stating that they would listen to and evaluate the quality of an innovative radio program. Participants used headphones and the experimenter (blind to the study conditions) ran the audio recording of the ostensible radio program.

Participants listened to one of five different versions of the radio program, which consisted of an interview with an individual who was in need of help. The radio recording began with the program host introducing Cátia, allegedly an ISCTE-IUL student whose parents died in a car accident the week before the recording date, and who had been given temporary guardianship of her two younger siblings. Cátia was said to be desperately in need of money to support herself and her siblings to the point that she did not have money for food. She also hoped to finish her undergraduate degree that year. This information was followed by an excerpt of the interview, in which she emphasized the extent of her need, so that she could graduate and get a decent job that would allow her to support her siblings, thereby avoid having to put them in an institution. In the no-aggression information condition, this was the extent of the information given. In the aggression information conditions she additionally reported an episode that occurred at her workplace where she behaved aggressively toward a colleague.

Independent variables

In all the aggression conditions, Cátia explained that she had a colleague who usually gave her paperwork when she was too busy, and vice-versa. One day the colleague asked Cátia to prepare some letters for clients (e.g., personalizing the contents of the letter for each client, printing the envelopes and stuffing the envelopes with the correct letter), which she spent a whole afternoon doing. When Cátia took the envelopes to her colleague the colleague explained that because she was distracted, she had not told Cátia that their boss had changed his mind, and wanted the letters sent by email.

The aggressive behavior described by the target included a manipulation of information elements that have been shown to prime the two punishment motives (Carlsmith et al. 2002; Darley et al. 2000). Specifically, we manipulated the following variables:

Harm magnitude (seriousness of the aggression) was manipulated to vary the just deserts motive through the nature of the aggressive behavior: a slap (high harm) or a refusal to share needed files (low harm). These aggressive behaviors were chosen based on a pretest conducted by us with 34 undergraduate students in which participants responded to the "Cuestionario de Actitudes Morales y Agresión" (CAMA; Andreu, et al., 2006). The CAMA classifies the degree of justification (i.e., the seriousness) of eight aggressive behaviors depending on their motivation and context. For the high harm magnitude conditions, we chose a hitting behavior, used as a punishment. Killing and torturing were considered in this pretest, irrespectively of the motive and context of occurrence, the least acceptable, but were not adequate for our study. Hitting as punishment was rated, immediately after killing and torturing, as less acceptable in the original Spanish study. We also thought about using hitting to preserve reputation, but it is possible that participants might forgive the target because of thinking that keeping a good reputation in the workplace would be a right to hold on to; especially in

45

the scenario we convey, in which the target could be seen as not able to afford losing her job due to getting a worse or bad reputation. In our view, being ironic would also not sound antisocial enough to be worthy of punishment. Thus, hindering was selected for the low harm magnitude conditions because it was considered, after being ironic, the most acceptable aggressive behavior.

In the *high* harm conditions, Cátia stated that in response she slapped her colleague on the face; whereas in the *low* harm conditions her colleague asked Cátia for the Word files she created for each client, which would spare her colleague a lot of time, and Cátia refuses to give them (no mention of slapping the colleague was made).

Target's past offense history (how often the offense occurred) was manipulated to vary the incapacitation motive. In the high past offense history condition information was included indicating that Cátia had problems with her team at her workplace, that she is vengeful, and that the situation described (whether involving the slap or the refusal to share files) is a usual problem. In the low history conditions this information was not included. Thus, in the aggression information conditions the recording continued with the host mentioning that he knew Cátia had "recently been through an unfortunate episode" (*low* history conditions), or "been having problems with her team" (*high* history conditions) at her workplace, and asking whether she wanted to talk about it. Cátia responded that she had no problem sharing it.

In the *high* history conditions Cátia began the description of the incident by saying she was a little vengeful. After the description, the host asked whether the reported situation was usual, and Cátia responded affirmatively.

More punishment and less helping when the offense had occurred multiple times (high history) rather than once (low history), and helping less in the low harm/high history than in the no aggression condition, would suggest an incapacitation motive for the punishment because, as previously stated, a history of aggressive behavior has been found to increase punishing with such motive and it was not present in the other conditions (low history and no-aggression conditions); whereas the same result when the magnitude of the harm caused was higher would suggest a just deserts motive, because the harm magnitude has been shown to be relevant for punishing with a just deserts motive, and the other conditions had a low harm or no aggression (low harm magnitude or no-aggression conditions).

Finally, in all the conditions, the host thanked Cátia for sharing her situation and ended the program by saying that the story was being broadcast with the hope that some of the listeners would help out, providing a phone number for contributions. A transcription of the complete recordings can be found in the attachments.

When the recording finished, instructions appeared on the screen asking participants to call the experimenter, who gave a questionnaire to the participant (see below). When she finished responding, the experimenter pretended, from his side of a folding screen, to be looking for a second

questionnaire. The experimenter then approached the participant with an envelope (containing a letter with the measure of willingness to help – see below) and told her that he had just realized he was out of the second questionnaire and would need to print more – but meanwhile the participant could read a letter that was supposed to be given when the study was finished. The experimenter then left the room. After 5 minutes, the experimenter returned, and told the participant to inform him when she was finished reading the letter so that he could give her the second questionnaire. Unless otherwise stated all items were rated in a 7 point scale from 1 to 7, with relevant anchors.

Dependent measures

The first questionnaire consisted of two items focusing on irrelevant aspects of the scenarios (e.g., Cátia's level of concern for her siblings) and a third item asking how severe a punishment Cátia should receive for her improper behavior ("Na entrevista, Cátia Crespo relata uma situação ocorrida no seu trabalho que culminou com um comportamento impróprio, que a Cátia não deveria ter tido. Quão severa considera que uma punição deveria ser dada à Cátia, pelo seu comportamento impróprio?"; in English: "In the interview, Cátia Crespo reports a situation that occurred in her work which culminated with an improper/inappropriate behavior that Cátia should not have done. How severe do you think a punishment should be given to Cátia, because of her improper behavior?"). The scale for answering ranged from 1 (*not severe/none*) to 7 (*extremely severe*). The second questionnaire asked participants to rate the radio program on eight characteristics (e.g., informative), on a 5-point scale from 1 (*not at all*) to 5 (*absolutely*); plus, an item asking participants to rate the seriousness of Cátia's need situation from 1 (*not severe at all*) to 7 (*very severe*). To check harm magnitude participants were asked how serious they considered Cátia's improper behavior, from 1 (*not serious at all*) to 7 (*very serious*); to check past offense history an item asked to what extent participants considered Cátia to be likely to behave improperly in the future, at her workplace, on a scale from 1 (*no likely at all*) to 7 (*very likely*). These questions included "Cátia reported no improper behavior" as an option. All participants received these questionnaires, but responses given by participants in the no-aggression condition on the item asking to assign a punishment severity to Cátia and on the manipulation check items were not analyzed because it did not make sense.

In the letter, ostensibly written by the professor responsible for the study, participants were asked to volunteer time (performing simple tasks that could be done at the student's home) for a fundraising campaign to help the target. The number of hours they indicated they were willing to volunteer was the help dependent measure.

The letter (a) thanked the participant for improving the quality of radio in the country; (b) reinforced that the situation described was real and representative of many other situations where people's lives have been torn apart and they must strive to recover; (c) mentioned that the university

47

and the radio station would like to help Cátia in return for her gesture sharing the story by helping her contact institutions and individuals (through phone calls and formal letters) to solicit financial contributions. It stated that the urgency of the situation led the professor to request the help of students in the campaign. Volunteers would help by preparing letters to mail (tasks could be done at the student's home). Then participants were asked to indicate the number of hours they would be willing to volunteer. Additionally, it stated that the process would take a month, so participants could distribute the hours they indicated as they pleased throughout the month.

Participants were asked to return the letter to the envelope (regardless of their decision), seal it, and place it in a box (already containing two sealed envelopes). After participants responded to the second questionnaire, the experimenter stopped the SCR recording and removed the electrodes. Participants washed their hands, and then the experimenter, with a notebook in hand, asked them about the program (what they enjoyed most and least, how competent the host seemed, whether he conducted the program and interview well) and probed for suspicion. When participants expressed suspicion, he tried to understand when they began suspecting and/or why (without revealing the deception). When the experimenter was sure about the effectiveness of the deception, he carefully started the oral debriefing, gave the written debriefing, thanked participants, and dismissed them.

The written debriefing complemented the previous oral debriefing, to reduce the chances that participants would discuss the experiment with their colleagues. The document reiterated the fact that the people and the situation were fictitious, and briefly explained, in simple terms, the goals of the study and the reasons for the deception. It also emphasized that communicating the true goals and procedures used in the study to other potential participants would produce invalid data and make all the sessions useless. The debriefing included a place to check that they would not discuss the study with other potential participants, and a signature line. The form also indicated that their signature authorized the use of the data for research purposes.

Results

Manipulation checks

To check whether the harm magnitude manipulation was understood by participants as expected, we submitted the seriousness of the improper behavior variable to a 2 (harm magnitude: high vs. low) X 2 (history: high vs. low) Analysis of Variance (ANOVA). The condition of no aggression was not included in the analysis. Results showed a main effect of harm magnitude, $F(1, 33) = 5.71$, $p = .02$, $\eta_p^2 = .15$, indicating that participants in high harm magnitude conditions considered the slap more serious ($M = 4.7$, $SD = 1.59$) than those in the low harm magnitude conditions ($M = 3.31$, $SD = 1.55$). No other significant results were found ($ps > .1$). The likelihood of behaving improperly in the future variable

48

was also subjected to the same two-way ANOVA, producing a significant main effect of history, $F(1,$ 33) = 7.01, p = .01, η_p^2 = .18, in the expected direction: participants in the high history conditions considered the target to be more likely to recidivate (M = 5.48, SD = 0.84) than those in the low history conditions (M = 4.13, SD = 1.6). No other significant results were found (ps > .1).

Hypothesis testing

We submitted each of the following dependent variables to the same two-way ANOVAs with the harm magnitude (high vs. low) and the history (high vs. low) as between-subject variables. The analyses including the no-aggression condition responses were one-way ANOVAs with experimental condition (high harm/high history vs. high harm/low history vs. low harm/high history vs. low harm/low history vs. no-aggression) as a between-subjects variable. Results are described below.

Punishment motives

We predicted that we would obtain, on punishment severity, a main effect of harm magnitude (H1.1.), or a main effect of history (H2.1.), or main effects of both harm magnitude and history (H3.1.). The two-way ANOVA produced no statistically significant effects (Fs < 2, ps > .1). Thus, we were not able to confirm hypotheses H1.1., H2.1., and H3.1..

Perceived seriousness of need situation

Results of the two-way ANOVA on perceptions of the seriousness of Cátia's need showed a significant main effect of harm magnitude, $F(1, 33)$ = 7.63, p = .009, η^2_p = .19, indicating that participants in high harm magnitude conditions (hitting) considered the target's need situation to be more serious (M = 6.65, SD = 0.49) than those in the low harm magnitude (hindering) conditions (M = 5.91, SD = 0.92). The one-way ANOVA revealed no significant differences between conditions (ps > .05). Contrasts showed that participants in both the high harm/high history (M = 6.60, SD = 0.52) and those in the high harm/low history (M = 6.70, SD = 0.48) perceived the need situation of the target as more serious than those in the low harm/high history condition (M = 5.80, SD = 0.50), $t(44)$ = 2.27, p = .03, Cohen's d = 1.57. Finally, participants in the high harm/low history condition perceived the need situation as more serious than those in the low harm/low history (M = 6.07, SD = 1.03), $t(44)$ = 2.08, p = .04, Cohen's d = 0.78. No other significant differences were detected, Fs < 1, ps > .1 (Figure 5.1.1.).

Figure 5.1.1.

Perceived Seriousness of the Target's Need Situation as a Function of Harm and History (Pre-test)
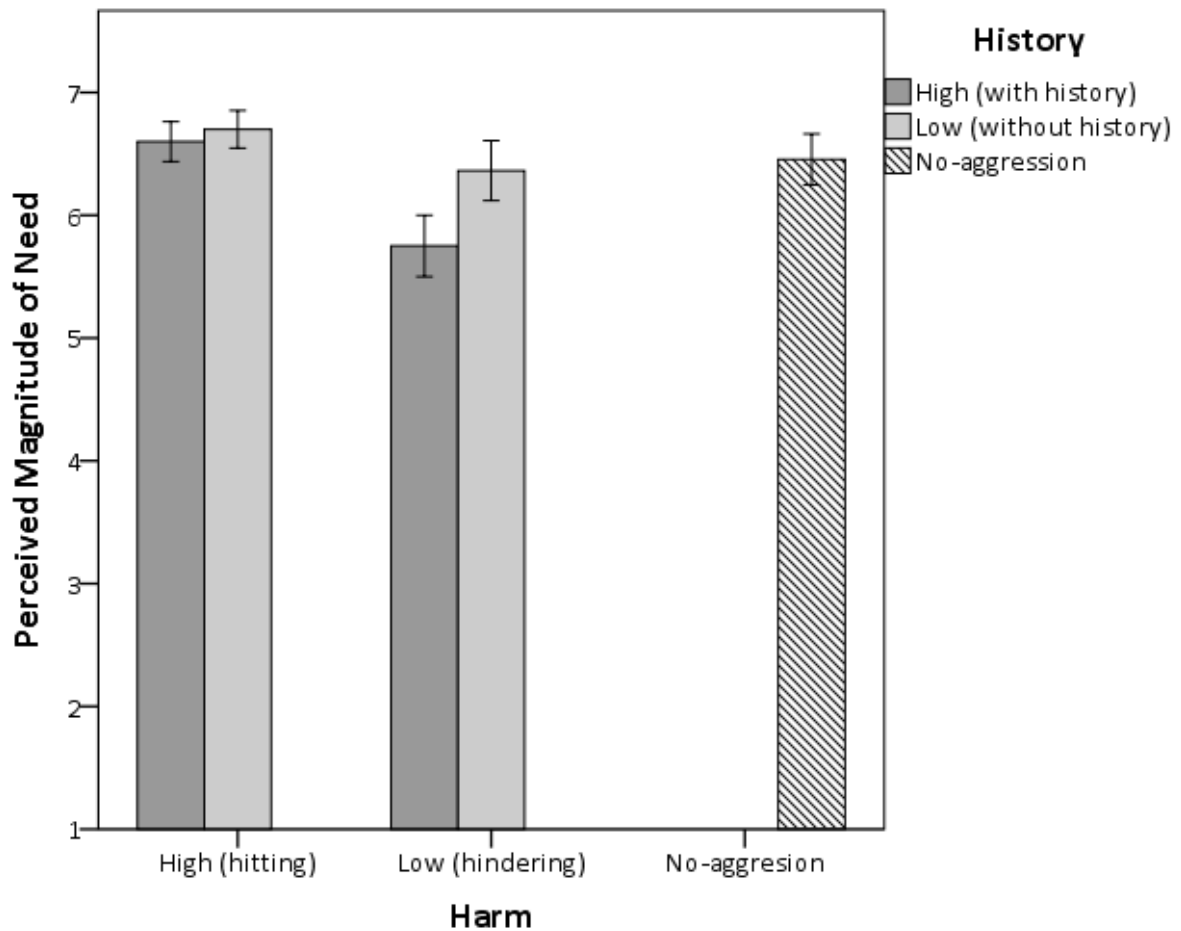
Note. Error bars represent standard errors.

Volunteered hours (help measure)

We expected to obtain a main effect of harm magnitude, with participants in high harm conditions helping less than those in low harm conditions (H1.2.), a main effect of the target's past offense history with participants in the high history conditions helping less than those in the low history conditions (H2.2.), or both (H3.2.). We also expected participants in the high harm/low history condition to help less than those in the no-aggression condition (H1.3.). Likewise, we expected participants in the low harm/high history condition to help less than those in the no-aggression condition (H2.3.).

Results of the two-way ANOVA showed a significant interaction between harm magnitude and history, $F(1, 31) = 6.81$, $p = .01$, $\eta^2_p = .18$. An analysis of simple effects showed that participants in the high harm/low history condition volunteered more time ($M = 10.80$, $SD = 10.93$; hitting once) than those in both the high harm/high history condition ($M = 0.56$, $SD = 1.13$; hitting multiple times), $F(1,31)$

= 13.56, p = .001, η^2_p = .30, and those in the low harm/low history condition (M = 0.83, SD = 1.64), F(1,31) = 14.78, p = .001, η^2_p = .32. No significant differences between the low harm conditions were found (p = .69).

The one-way ANOVA revealed significant differences between conditions, F(4, 42) = 5.67, p = .001, η^2_p = .35. Contrasts showed that participants in the high harm/low history condition volunteered more hours (M = 10.80, SD = 10.93; hitting once) than those in the no-aggression (M = 2.83, SD = 3.97), t(42) = -3.33, p = .002, Cohen's d = 0.97, and also than all the remaining aggression conditions, t(44) = 2.69, p = .01, Cohen's d = 1.16. No other contrasts were significant, Fs < 0.2, ps > .10 (see Figure 5.1.2).

Figure 5.1.2.

Number of Volunteered Hours for the Future Fundraising Campaign as a Function of Harm and History (Pre-test)
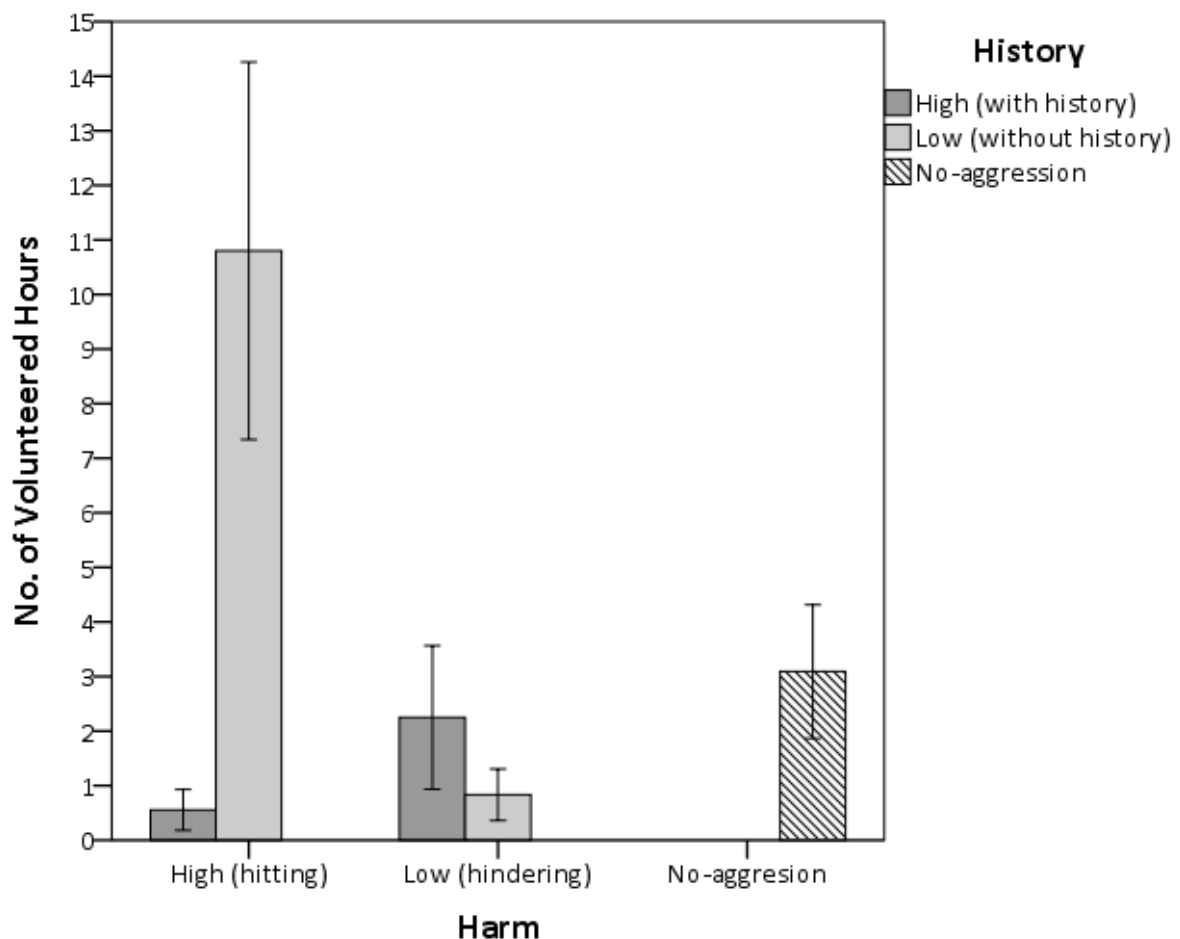
Note. Error bars represent standard errors.

We used the dichotomized variable of the hours volunteered (1 = volunteered time and 0 = volunteered no time) to compare the proportion of participants who volunteered and did not volunteer time. We found that in the high harm conditions the proportion of participants who withheld help (by volunteering no time at all) was larger (n = 10) than the proportion of participants who helped by volunteering time (n = 9), χ^2(1, *n* = 19) = 4.38, *p* = .04. Likewise, in the low history conditions, the proportion of participants who withheld help was larger (n = 12) than the proportion of participants who helped by volunteering time (n = 10), χ^2(1, n = 22) = 3.50, *p* = .06.

To compare the proportion of participants who volunteered and did not volunteer time using a non-parametric test we conducted a Kruskal-Wallis test. No significant differences were found in the proportions of participants who volunteered and did not volunteer between conditions (*p*s > .60). Table 5.1. presents the means and standard deviations for all the dependent variables.

Table 5.1.

Means and Standard Deviations for Punishment Severity, Perceived Seriousness of Need Situation, and Volunteered Hours, by Condition (Pre-test)

Condition	Dependent measure		
	Punishment severity	Perceived seriousness of need	Volunteered hours
No aggression (n = 12)	-	6.50 (0.67)[a]	2.83 (3.97)[b]
Low Harm/Low History (n = 13)	3.38 (1.39)	6.07 (1.03)[c]	0.83 (1.64)[b]
Low Harm/High History (n = 4)	3.00 (1.41)	5.75 (0.50)[b]	2.25 (2.63)[b]
High Harm/Low History (n = 10)	4.00 (1.89)	6.70 (0.48)[a]	10.80 (10.93)[a]
High Harm/High History (n = 10)	3.89 (1.05)	6.60 (0.52)[a]	0.56 (1.13)[b]

Note. Means that share a superscript do not differ significantly at *p* = .05, using paired *t*-tests.

Discussion

With the pretest we wanted to test the paradigm used frequently in the US in a new culture and language and extend the paradigm to include additional information about the target, Cátia. Manipulation checks supported our expectations that hitting would be seen as causing a higher magnitude of harm than hindering work, and that when Cátia described herself as having behaved in the negative way multiple times, participants would rate future similar behavior as more likely. However, neither higher harm nor higher history consistently increased punishment (H1.1. and H2.1.,

52

respectively; though those in the high harm conditions tended to punish the target more severely than those in the low harm conditions) nor decreased volunteered time to help the target (H1.2., H1.3., H2.2, and H2.3.). In fact, volunteered time was greatest in the case when she had hit a colleague once (with a medium effect size). We found no evidence for confirmation of hypotheses H1, H2, and H3.

It is possible that participants may have taken cues about the seriousness of the target's need situation from the information about her behavior. Participants in the hitting conditions may have considered the slap so serious, or unacceptable, that only a very distressed person could have committed it. Consistent with this explanation, previous research has shown that an important predictor of just deserts punishment is the absence of mitigating circumstances (Carlsmith, 2006), such as contextual factors (Carlsmith et al., 2002). In this case it is possible that participants attributed more emotional distress to the target when she slapped her colleague, and thus did not withhold help. Participants may have also increased their help, thinking that if they helped relieve Cátia's stress, she would be less likely to behave in a similar way in the future. The slap could have also increased the participants' perceived need of Cátia, leading to more compassion felt for her, and consequently to more help. Compassion is defined as a "a strong feeling of sympathy with another person's feelings of sorrow or distress, usually involving a desire to help or comfort that person" (VandenBos, 2015).

These possibilities led us to consider how intentional participants saw the target's behavior and how responsible for it they held her. We reasoned that participants in the just deserts priming condition (high harm/low history) may have considered the slap an uncontrolled or unintentional behavior, which the target should not be considered responsible for, because of emotional distress of having lost her parents in a car accident. Accordingly, in the full study we added items measuring the target's emotional stability, intention to harm her colleague, responsibility for her act and compassion felt for her.

5.2. STUDY 1

Aims and Hypotheses

The Pretest we conducted was designed to understand whether our materials, procedures, and measures would effectively test our hypotheses. The results and our interpretations led us to the addition of measures in Study 1. The materials, procedures, and hypotheses of Study 1 were identical to those of the Pretest.

<center>**Method**</center>

Participants

One hundred and seven undergraduate students (aged 18-32; *M* = 21.10, *SD* = 2.45) participated in the study in exchange for course credit. As in the Pretest, participants were asked to listen to and evaluate a bogus pilot radio program. They were randomly assigned to one condition of a 2 (harm magnitude: high vs. low) X 2 (history: high vs. low) or a control condition (no-aggression information) between subjects design. Only female subjects participated.

In the Pretest and Study 1, ten participants were excluded from the sample due to suspicion about the veracity of the recordings (four from the high harm magnitude/high history conditions; three from the low harm magnitude/high history conditions; one from the low harm magnitude/low history condition; two from the no-aggression conditions).

We conducted power analysis using G*Power 3.1 (Faul, et al., 2009) to compute the required sample size for an a-priori ANOVA with a power of 80%, a medium effect size (f = 0.25), and an error probability of 5% significance. The ideal sample size would be 196 participants (39 per condition). Unfortunately, we were not able to reach that number of participants but collected data from as many participants as possible in the time available.

Measures and procedure

The procedure for Study 1 was the same as that of the Pretest with some changes to the items in some of the questionnaires. In the first questionnaire, the item asking how severe a punishment Cátia should receive for her improper behavior, as well as the manipulation check items, were only included for participants who received the aggression information; participants in the no-aggression condition only responded to two items about irrelevant aspects of the scenarios. In order to better understand participants responses to the target in different conditions, two items were added to the second questionnaire in all conditions, plus another two to the second questionnaire in the aggression information conditions. The two items that were added to the second questionnaire in all conditions were: a) an item asking participants how emotionally stable they considered Cátia, on a scale from 1 (*not emotionally stable at all*) to 7 (*very emotionally stable*), placed before the item measuring the seriousness of her need; b) after the item measuring the seriousness of her need, an item asking participants how much compassion they felt for Cátia (1 = not compassionate at all; 7 = extremely compassionate).

In addition to these, the second questionnaire for all the aggression information conditions also included an item asking how intentional participants considered Cátia's improper behavior (1 = not intentional at all; 7 = totally intentional), and an item asking how responsible for that behavior they

54

considered her (1 = not responsible at all; 7 = totally responsible). Because asking participants in the aggression conditions to assign a punishment severity could influence their volunteered time to help, the order of delivery of the second questionnaire and the letter with the help measure was counterbalanced[4].

Results

Manipulation checks

The harm magnitude manipulation was checked by submitting the seriousness of the improper behavior variable to a 2 (harm magnitude: high vs. low) X 2 (history: high vs. low) ANOVA. Results showed a main effect of harm magnitude, $F(1, 83) = 14.16$, $p < .001$, $\eta^2_p = .15$, indicating that participants in high harm magnitude conditions considered the slap more serious ($M = 5.08$, $SD = 1.42$) than those in the low harm magnitude conditions considered the hindering behavior ($M = 4.05$, $SD = 1.18$). The ANOVA also resulted in a main effect of history, $F(1, 83) = 6.81$, $p < .04$, $\eta^2_p = .05$, showing that participants who learned of more than one occurrence of aggression (high history conditions) considered that more serious ($M = 4.76$, $SD = 1.32$) than those who learned of only a single occurrence of aggression (low history conditions; $M = 4.29$, $SD = 1.44$). The interaction was non-significant ($p = .33$). We also subjected the likelihood of behaving improperly in the future variable to the same ANOVA, which produced a significant main effect of the history, $F(1, 82) = 13.68$, $p < .001$, $\eta^2_p = .14.$, showing that participants in the high history conditions considered the target to be more likely to recidivate ($M = 4.81$, $SD = 1.40$) than those in the low history conditions ($M = 3.70$, $SD = 1.38$). No other significant results were found (ps > .05).

Like in the pre-test, we submitted each of the following dependent variables to the same: (a) two-way ANOVA with the harm magnitude (high vs. low) and the history (high vs. low) as between-subject variables; and (b) one-way ANOVA with the experimental condition (high harm/high history vs. high harm/low history vs. low harm/high history vs. low harm/low history vs. no-aggression) as a between-subjects variable. Results are described below.

---

[4] We tested whether the order of receiving the second questionnaire and the letter influenced results of volunteered time to help by running a three-way ANOVA with a 2 (harm magnitude: high vs. low) X 2 (past offense history: high vs. low) X 2 (order: questionnaire first vs. letter first) design. No statistically significant results were produced, therefore we collapse across order for all other reported results.

Hypothesis testing

Punishment motives

We ran the described two-way ANOVA in order to check the activation of the punishment motives, with punishment severity as the dependent variable. We predicted that we would obtain, on punishment severity, a main effect of harm magnitude (H1.1.), or a main effect of history (H2.1.), or main effects of both harm magnitude and history (H3.1.).

Results showed a reliable main effect of harm magnitude, $F(1, 83) = 4.62$, $p = .03$, $\eta^2_p = .05$, indicating that participants in the high harm magnitude conditions (hitting behavior) assigned a more severe punishment ($M = 4.05$, $SD = 1.36$) than those in the low (hindering; $M = 3.44$, $SD = 1.38$; confirming H1.1.). Similarly, there was a main effect of history, $F(1, 83) = 5.66$, $p = .02$, $\eta^2_p = .06$, indicating that participants in the high history conditions punished the target more severely ($M = 4.09$, $SD = 1.16$) than those in the low ($M = 3.4$, $SD = 1.55$; confirming H2.1.). Since the ANOVA resulted in the main effects of both harm magnitude and history on punishment severity, H.3.1. was also confirmed. The interaction was non-significant ($p = .32$).

Perceived seriousness of need situation

Both the two- and one-way ANOVAs produced no statistically significant results, Fs < 2, ps > .1.

Volunteered hours (help measure)

Our predictions were that participants in the high harm conditions would volunteer less time than participants in the low harm conditions (H1.2.), and that participants in the high harm/low history condition would volunteer less time than and those in a no-aggression condition (H1.3.). We also expected participants in the low harm/high history condition to help less than those in the no-aggression condition (H2.3.), and that participants in the high history conditions would volunteer less time than participants in both low history conditions (H2.2.), or both (H3.2.).

Results of the two-way ANOVA revealed no significant main effects, failing to support our hypotheses. However, a significant interaction between harm magnitude and history was found, $F(1, 73) = 5.33$, $p = .02$, $\eta^2_p = .07$. An analysis of simple effects showed that participants in the high harm/high history condition volunteered more time ($M = 3.95$, $SD = 6.40$) than those in low harm/high history condition ($M = 0.29$, $SD = 0.72$), $F(1,73) = 7.64$, $p = .007$, $\eta^2_p = .10$. No significant differences were found between the low harm/high history condition and the low harm/low history condition ($M = 2.71$, $SD = 4.75$), $p = .08$, nor between the two high harm conditions ($p = .14$), nor between the two low history conditions ($p = .59$).

Results of the one-way ANOVA showed no significant differences between conditions. However, subsequent contrasts showed that participants in the low harm/high history condition (repeated hindering – incapacitation priming) volunteered significantly less time (M = 0.29, SD = 0.72) than participants in the no-aggression information condition (M = 3.06, SD = 4.14), $t(17.88)$ = 2.81, p = .01, Cohen's d = 0.93, suggesting our prediction about withholding help with an incapacitation motive. Participants in the low harm/high history (incapacitation priming) condition also volunteered less time than those in both the low harm/low history (M = 2.71, SD = 4.65), $t(16.62)$ = -2.13, p = .05, Cohen's d = 0.71, and the high harm/low history (just deserts priming; M = 1.95, SD = 3.12), $t(20.91)$ = 2.39, p = .03, Cohen's d = 0.73. The difference in volunteered time between the high harm/low history (just deserts priming condition) and the no-aggression condition was not significant, thus providing no support for help withheld with a just deserts motive (Figure 5.2.), but partially replicating the results of the pretest.

Finally, we created a dichotomized variable of the hours volunteered (1 = volunteered time and 0 = volunteered no time) to compare the proportions of participants who volunteered and did not volunteer time. No significant differences were found, ps > .1.

Figure 5.2.

Number of Volunteered Hours as a Function of Harm and History (Study 1)
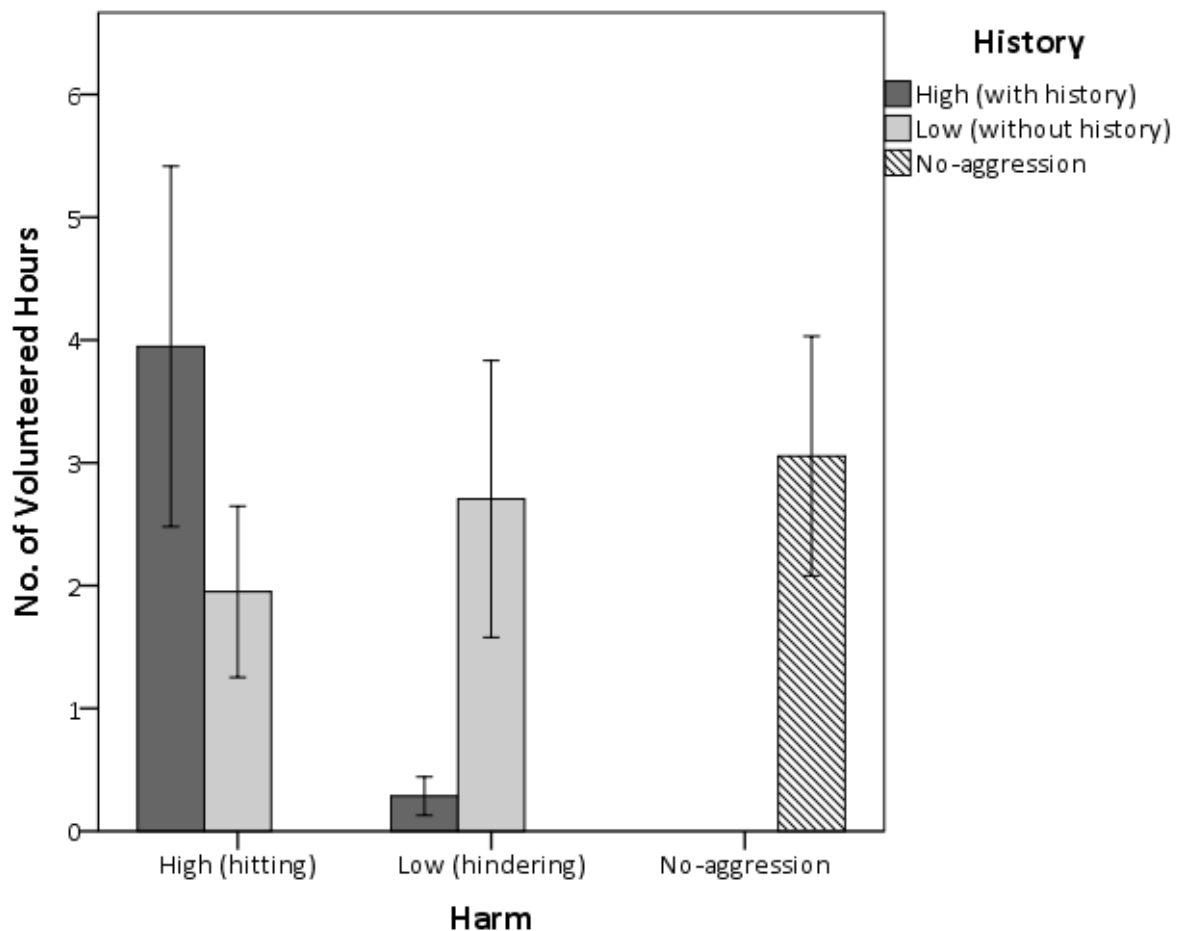
Note. Error bars represent standard errors.

Emotional stability, intentionality, responsibility and compassion

To test differences in interpretations of the target across condition, emotional stability, intentionality, responsibility and compassion were individually submitted to the same ANOVAs. Results of the two-way ANOVA on the intentionality showed a significant main effect of the harm magnitude, $F(1, 83) = 11.85$, $p = .001$, $\eta^2_p = .13$, indicating that participants in the high harm conditions considered the (hitting) behavior less intentional ($M = 3.15$, $SD = 1.64$) than those in the low harm (hindering) conditions ($M = 4.38$, $SD = 1.89$). There was also a significant main effect of history, $F(1, 83) = 11.85$, $p = .001$, $\eta^2_p = .13$, such that participants in the high history conditions considered the aggressive behavior more intentional ($M = 4.38$, $SD = 1.76$) than those in the low history conditions ($M = 3.15$, $SD = 1.78$). The interaction was not significant ($p = .08$).

Results of the two-way ANOVA on responsibility resulted in a significant main effect of the history, $F(1, 81) = 7.61$, $p = .007$, $\eta^2_p = .09$, such that participants in the high history conditions attributed more responsibility to the target for her aggressive behaviors ($M = 5.09$, $SD = 1.61$) than their low history counterparts ($M = 4.14$, $SD = 1.72$). This effect was qualified by a significant interaction with harm magnitude, $F(1, 81) = 10.25$, $p = .002$, $\eta^2_p = .11$. An analysis of simple effects showed that participants in the low harm/low history condition assigned less responsibility to the target ($M = 3.33$, $SD = 0.97$) than participants in both the high harm/low history ($M = 4.95$, $SD = 1.88$), $F(1,81) = 10.27$, $p = .002$, $\eta^2_p = .11$, and those in the low harm/high history conditions ($M = 5.38$, $SD = 1.49$), $F(1,81) = 18.10$, $p < .001$, $\eta^2_p = .18$. No significant differences between the high harm conditions ($p = .78$) were found, nor between the high history conditions ($p = .22$).

Results of the one-way ANOVA showed significant differences between conditions, $F(3, 81) = 6.38$, $p = .001$, $\eta^2_p = .19$. Contrasts showed that participants in the low harm/low history condition attributed the lowest responsibility, when compared with all the remaining conditions, $F(1, 81) = 14.83$, $p < .001$, $\eta^2_p = .16$. On the other hand, participants in the low harm/high history (repeated hindering - incapacitation priming) attributed the highest responsibility, also when comparing with the remaining conditions, $F(1, 81) = 8.34$, $p = .005$, $\eta^2_p = .09$. Intentionality and responsibility were positively and significantly correlated, $r = .31$, $p = .004$.

Ratings of emotional stability were low on average ($M = 3.24$, $SD = 1.37$), and both the one-way and the two-way ANOVAs showed no significant differences between conditions. Similarly, the reported compassion felt for the target was high in all conditions ($M = 5.98$, $SD = 1.03$), and no significant differences were found using both ANOVAs. Contrasts run on compassion revealed, however, that participants in the high harm/low history (just deserts priming) reported more compassion ($M = 6.28$, $SD = 0.669$), than participants in the remaining aggression conditions, $F(1, 97) = 4.13$, $p = .045$, $\eta^2_p = .04$, but not higher than the compassion felt by those in the no-aggression condition ($M = 6.35$, $SD = 6.71$, $p > .10$).

See Table 5.2. for means and standard deviations of all the dependent variables.

Table 5.2.

Means and Standard Deviations for Punishment Severity, Perceived seriousness of need situation, Volunteered hours, Intentionality, Responsibility, Emotional Stability, and Compassion by Condition (Study 1)

Condition	Dependent Measure						
	Punishment severity	Perceived seriousness of need	Volunteered hours	Intentionality	Responsibility	Stability	Compassion
Low Harm / Low History – hindering once (n = 20)	2.95 (1.47) [b]	6.45 (0.69) [a]	2.71 (4.75) [a]	3.45 (1.99) [b]	3.33 (0.97) [c]	3.00 (1.39) [a]	5.75 (1.45) [b]
Low Harm / High History – repeated hindering (n = 26)	3.92 (1.16) [a]	6.50 (0.58) [a]	0.29 (0.72) [b]	5.31 (1.38) [a]	5.38 (1.49) [a]	3.19 (1.49) [a]	5.75 (0.79) [b]
High Harm / Low History – hitting once (n = 21)	3.86 (1.53) [a]	6.70 (0.47) [a]	1.95 (3.12) [a]	2.86 (1.59) [b]	4.95 (1.88) [b]	3.14 (.91) [a]	6.28 (0.67) [a]
High Harm / High History – repeated hitting (n = 20)	4.25 (1.16) [a]	6.45 (0.69) [a]	3.95 (6.40) [a]	3.45 (1.67) [b]	4.80 (1.74) [b]	2.95 (1.13) [a]	5.85 (1.27) [b]
No-aggression (n = 20)	-	6.50 (0.61) [a]	3.06 (4.14) [a]	-	-	3.90 (1.71) [a]	6.35 (0.67) [a]

Note. Means that share a superscript do not differ significantly at *p* = .05, using paired *t*-tests.

Discussion

Our original hypotheses were that participants would withhold help as a form of punishment when the target behaved aggressively in situations where relevant information for the just deserts (high harm/low history; H1.2. and H1.3.) and incapacitation (low harm/high history; H2.2. and H2.3.) punishment motives were present. Given the Pretest results, however, the alternative hypothesis that only the incapacitation motive would lead to withholding of help because of feelings of compassion for the target and thus a lack of assignment of intentionality and responsibility for the aggressive behavior was considered as a possible alternate result.

Differently from the results obtained in the Pretest, in this study we found that participants activated both punishment motives in assigning the punishment severity (H1.1., H2.1., and H3.1.).

We also found that participants in this study expressed less willingness to help when the target repeatedly acted aggressively toward others by hindering their work (a low-seriousness aggression) (with moderate and large effect sizes), suggesting an incapacitation motive to stop the target from continuing the actions. However, the absence of main effects failed to support the hypothesis of help withholding with that motive.

On the other hand, a serious aggressive behavior performed one time (or more) did not reduce participants' willingness to help the target, suggesting a just deserts motive to restore justice was not activated as we expected it would be. We found no evidence for confirmation of hypotheses 1.2., 1.3., 3.2., and 3.3..

These results raise several questions. The first concerns the psychological process underlying punishing with an incapacitation goal: to what extent did participants believe that withholding help would prevent the target from future offenses (which would be necessary for an incapacitation goal to lead to less help) and how would it do this?

The incapacitation motive generally leads people to try to prevent contact between the offender and potential victims. In this experiment helping the target would improve her financial situation, reducing her need to work (thus isolating her from her work colleagues). Thus, if participants are thinking about the consequences of helping in such a rational manner, a decision not to help based on an incapacitation motive seems unlikely.

It is possible, however, that participants reasoned in a bounded or more heuristic fashion. They could have activated specific beliefs about having money and thought accordingly. For instance, participants could simply believe that the more money a person has, the more people that person is in contact with (or is able to contact or interact with), and then proceeded to a conditional: "if I help this aggressive person collect money, this aggressive person will be in contact with more people – and thus be aggressive to more people too (i.e., there will be more victims)". In this sense, helping the target might mean contributing to her opportunity to re-offend or something similar.

61

Another question pertains to the fact that in the high harm magnitude conditions participants did not help less than in the low. It is possible that they thought that, since the target's colleague was negligent in the first place (in not telling the target that their boss changed his mind), upsetting a person who is already suffering, the hitting behavior seemed just – thus making the target deserve help as much as if she hindered.

The result in the Pretest of a higher number of volunteered hours by participants in the high harm/low history condition (just deserts priming condition) led us to consider how intentional participants saw the behavior and how responsible for it they held the target. Although the effect on volunteer hours disappeared in Study 1, our instinct that the act of hitting someone was seen as less intentional than hindering their work was confirmed. Although results found in Study 1 were different from the pattern obtained in the pretest, both studies suggested that participants interpreted hindering as more intentional than hitting, and when done repeatedly, saw it as more deserving of punishment.

An alternative explanation for the lowest time volunteered in the low harm/high history condition dwells in another utilitarian theory of punishment for which the perpetrator's history of offenses is relevant: deterrence sentencing (see Utilitarian punishment theories: incapacitation and deterrence in Chapter 4). Having attributed the highest intentionality to Cátia's aggression in this condition, it seems reasonable that participants might have expected the target to make the association between the lack of help she received and her aggressive behaviors, hoping to dissuade her from harming in the future (thus having activated a deterrence motive). Our first study did not test this hypothesis.

Based on these questions, we designed a follow-up study using a different high harm magnitude behavior with an additional manipulation of deterrence motive.

5.3. STUDY 2

Aims and Hypotheses

To clarify the decision process underlying the help withheld by participants in the low harm magnitude/high history condition, we tested in this study the hypothesis that participants in that condition withheld help with a deterrence motive (H4). For testing the just deserts hypothesis (H1) we changed the serious aggression (high harm magnitude conditions) to one that would be perceived as intentional by participants: a threatening behavior. Finally, we included items in a third questionnaire measuring how much participants liked the target; agreed with sentences relating money, contacted people, and aggression victims; and concerning their past volunteering experience.

Since withholding help can only be a deterrence if the target knows that help is being withheld, participants with a deterrence motive for punishing Cátia should not withhold help if they believe that

Cátia will not know how much people help. Deterrence was thus manipulated by including motive inhibiting information (low target volunteer knowledge) or no additional information (high target volunteer knowledge) in the letter from the experimenter given to participants of both high history conditions. The deterrence inhibiting information consisted of a sentence saying that Cátia would not learn the number of hours volunteered individually or in total. In the high target volunteer knowledge conditions, no information about Cátia learning about the number of hours volunteered was included. If we see more helping when the sentence is present, it suggests participants are withholding help with a deterrence motive. In other words, participants are reasoning that Cátia will interpret the low amount of help as a punishment for her past aggressive behavior, which could deter her from acting aggressively in the future.

Thus, in Study 2 we maintained our predictions – custom hypotheses tested with planned contrasts – regarding:

Withholding help as a form of punishment with a just deserts motive (H1), i.e., we expected:

- o A main effect of harm magnitude on punishment severity, with participants in the high harm conditions punishing the target more severely than in the low conditions, indicating the activation of a just-deserts motive for punishing on the punishment severity item (H1.1.);
- o A main effect of harm magnitude on helping, with participants in the high harm conditions helping less than those in low harm conditions (H1.2.).

Withholding help as a form of punishment with an incapacitation motive (H2) (decision based on beliefs about having money), if we obtained:

- o A main effect of the target's past offense history on the punishment severity, with participants in the high history conditions punishing the target more severely than in the low conditions, indicating the activation of a utilitarian motive for punishing on the punishment severity item (H2.1.);
- o A main effect of the target's past offense history on helping, with participants in the high history conditions helping less than those in the low history conditions (H2.2.);
- o Participants in the low harm/high history/low target volunteer knowledge condition helping less than those both in a low harm/high history/high target volunteer knowledge and in the low harm/low history condition (H2.3.);
- o Participants in a low harm/high history/low target volunteer knowledge condition expressing agreement with both sentences "In general, the more money a person has, the more people that person contacts or interacts with" and "The more people an aggressive person contacts or interacts with, the more aggression victims that person causes." (H2.4.).

63

Additionally, we also predicted that participants would withhold help as a form of punishment with a deterrence motive (H4), if we obtained:

- o A main effect of the target's past offense history on the punishment severity, with participants in the high history conditions punishing the target more severely than in the low conditions, indicating the activation of a utilitarian motive for punishing in the punishment severity item (H2.1.);

- o A main effect of the target's past offense history on helping, with participants in the high history conditions helping less than those in the low history conditions (H2.2.);

- o Participants in a low harm/high history/high target volunteer knowledge condition volunteer less time than those both in the low harm/high history/low target volunteer knowledge and the low harm/low history conditions (H4.1).

Method

Participants

One hundred twenty-three undergraduate students participated in the study (aged 18-30; *M* = 19.96, *SD* = 0.23). Like in Study 1, participants were asked to listen to and evaluate a bogus pilot radio program. They were randomly assigned to one condition of a 2 (harm magnitude: high vs. low) X 2 (past offense history: high vs. low) X 2 (target volunteer knowledge: high vs. low) between-subjects design, with the harm magnitude and the history fully crossed, and target volunteer knowledge nested within the high history conditions. Target volunteer knowledge was used to manipulate the deterrence punishment motive. Only female subjects participated. No participants were excluded from the sample due to suspicion about the veridicality of the recordings. We conducted power analysis using G*Power 3.1 (Faul, et al., 2009) to compute the power of a post-hoc ANOVA using a medium effect size (f=.25) and an error probability of 5% significance. The achieved power was of 60%.

Measures and procedure

Study 2 replicated the procedure of Study 1, with some additional measures. Because in Study 1 there were no significant differences between the number of volunteered hours by participants in the no-aggression and the low harm/low history condition, we did not run the no-aggression condition in Study 2.

Participants listened to one of four versions of the radio program described in Study 1, all of which contained the description of an episode where Cátia behaved aggressively toward a colleague. The high harm magnitude manipulation was modified from Study 1 to describe a situation in which Cátia threatens rather than hits her colleague (Cátia says: "E eu respondi-lhe que ela agora tinha de entregar

64

as cartas mesmo assim, caso contrário eu contava ao nosso chefe de todas as vezes que ela saiu mais cedo do trabalho… e… que ela seria despedida depois disso.”).

The manipulation of the offense history was identical to the manipulation in Study 1. After responding to the second questionnaire, participants were given a third questionnaire containing the following measures: (a) an item measuring liking for Cátia, on a scale from 1 (*liked nothing*) to 7 (*liked very much*); b) an item asking how much participants agreed with the sentence “In general, the more money a person has, the more people that person contacts or interacts with”, on a scale from 1 (*totally disagree*) to 7 (*totally agree*); c) an item asking how much participants agreed with the sentence “The more people an aggressive person contacts or interacts with, the more aggression victims that person causes.”, also on a scale from 1 (*totally disagree*) to 7 (*totally agree*); and d) an item measuring participants previous volunteering activities (yes or no) and, if yes, how many hours they volunteer.

Results

Manipulation checks

The harm magnitude manipulation was checked by submitting the seriousness of the improper behavior variable to a 2 (harm magnitude: high vs. low) X 2 (history: high vs. low) ANOVA. Results showed a main effect of harm magnitude, $F(1, 118) = 4.14$, $p = .04$, $\eta^2{}_p = .3$, indicating that participants in high harm magnitude conditions considered the threat more serious ($M = 4.64$, $SD = 1.23$) than those in the low harm magnitude conditions considered the hindering behavior ($M = 4.15$, $SD = 1.28$). No other significant results were found (ps > .1). We also subjected the likelihood of behaving improperly in the future variable to the same ANOVA, which produced a significant main effect of the history, $F(1, 118) = 22.11$, $p < .001$, $\eta^2_p = .16$., showing that participants in the high history conditions considered the target to be more likely to recidivate ($M = 5.16$, $SD = 1.37$) than those in the low history conditions ($M = 3.83$, $SD = 1.63$). No other significant results were found (ps > .5).

Hypothesis testing

Each following dependent variable was submitted to the same: (a) two-way ANOVA with 2 (harm magnitude: high vs. low) X 3 (history: high with high target volunteer knowledge vs. high with low target volunteer knowledge vs. low) between subjects design; and (b) one-way ANOVA with the experimental condition (high harm / high history / high target volunteer knowledge vs. high harm / high history / low target volunteer knowledge vs. high harm / low history vs. low harm / high history / high target volunteer knowledge vs. low harm / high history / low target volunteer knowledge vs. low harm / low history) as a between-subjects variable. Results are described below.

Punishment motives

We predicted that we would obtain, on punishment severity, a main effect of harm magnitude (H1.1.), or a main effect of history (H2.1.). Both ANOVAs showed no significant results ($ps > .07$), so none of the hypotheses was confirmed. Further planned comparisons showed that participants in all three conditions enabling the punishment motives – high harm/low history condition (just deserts priming; $M = 3.60$, $SD = 0.88$), low harm/high history/high target volunteer knowledge condition (deterrence priming; $M = 3.60$, $SD = 1.35$), and low harm/high history/low target volunteer knowledge condition (incapacitation priming; $M = 3.37$, $SD = 1.34$) –, as well as those in the high harm/high history/high target volunteer knowledge condition ($M = 3.67$, $SD = 1.62$) did not assign more severe punishments to Cátia than did those in the low harm/low history condition ($M = 3.05$, $SD = 1.02$) (all $ps. > .1$). Thus, there was no indication of any punishment motive activated. Only participants in the high harm/high history/low target volunteer knowledge condition ($M = 4.00$, $SD = 1.30$) assigned more severe punishments to Cátia than those in the low harm/low history condition, $t(37.87) = -2.63$, $p = .01$, Cohen's $d = 0.81$. (Figure 5.3.1.).

Figure 5.3.1.

Punishment Severities as a Function of Harm, History, and Target Volunteer Knowledge (Study 2)
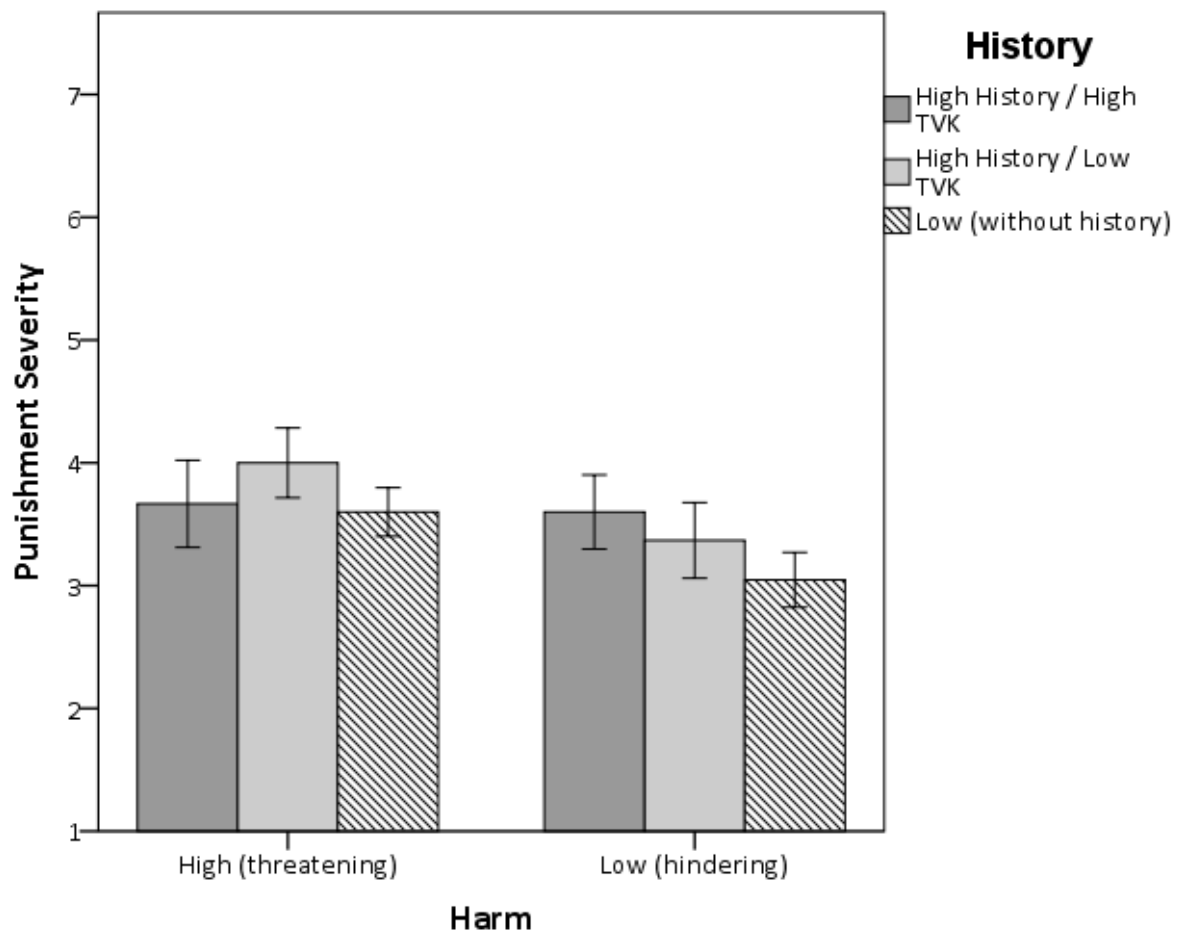
Note. TVK: Target Volunteer Knowledge. Error bars represent standard errors.

Perceived seriousness of need situation

Both the two-way and the one-way ANOVAs showed no significant differences between conditions (ps

> .1; Figure 5.3.2.; see Table 5.3. for means and standard deviations).

Figure 5.3.2.

Perceived Seriousness of the Target's Need Situation as a Function of Harm, History, and Target Volunteer Knowledge (Study 2)
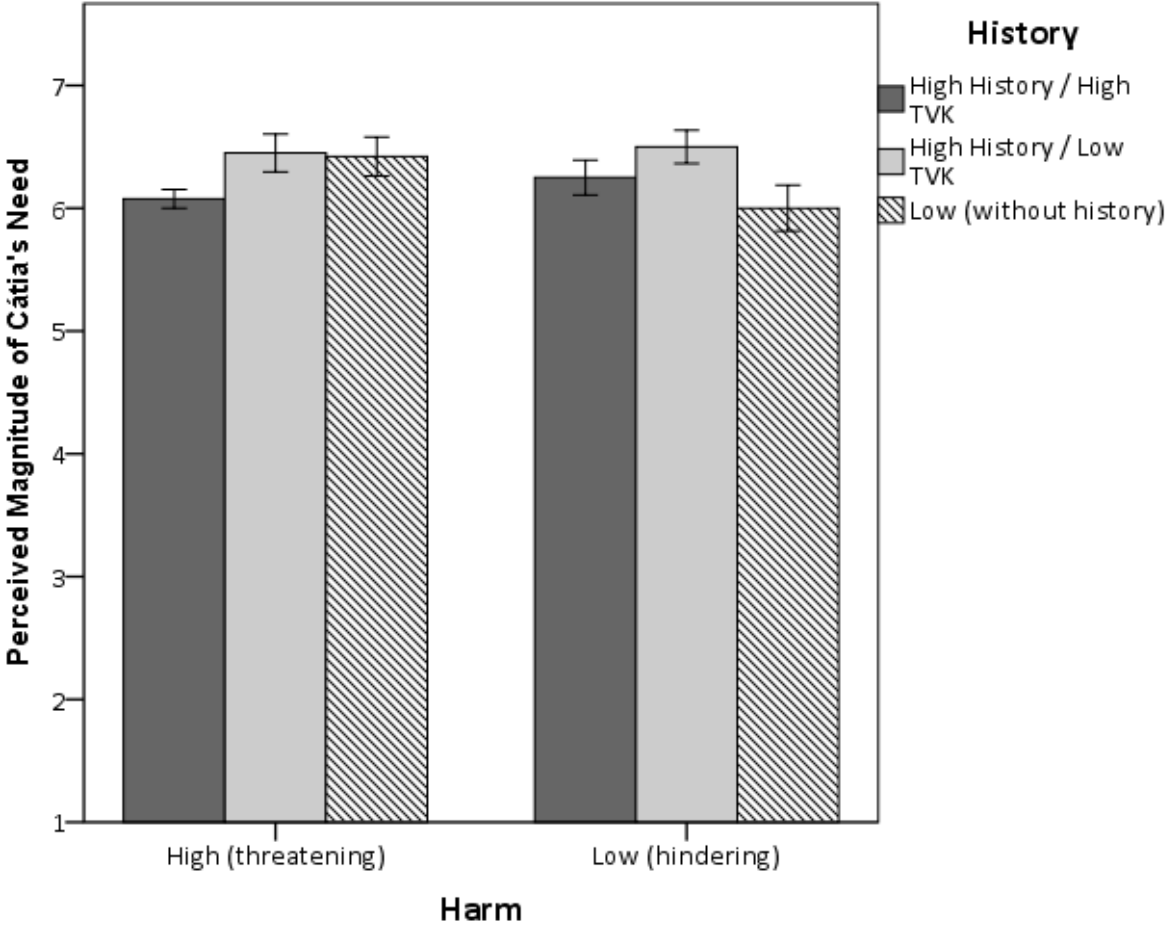
Note. TVK: Target Volunteer Knowledge. Error bars represent standard errors.

Volunteered hours (help measure)

In Study 2 we predicted a main effect of the target's past offense history on helping, with participants in the high history conditions helping less than those in the low history conditions (H2.2.). Participants in the low harm/high history/low target volunteer knowledge condition were expected to volunteer less time than those both in a low harm/high history/high target volunteer knowledge and in the low/low condition (H2.3.); and participants in a low harm/high history/high target volunteer knowledge condition were expected to volunteer less time than those both in the low harm/high history/low target volunteer knowledge and the low harm/low history conditions (H4.1). We also expected a main effect of harm magnitude on helping, with participants in the high harm conditions (just deserts priming, now using a threatening behavior) helping less than those in low harm conditions

(H1.2.). We tested these hypotheses, running planned comparisons of the volunteered hours in each condition.

Results of the two-way ANOVA showed a significant main effect of history, $F(2, 96) = 8.11$, $p = .001$, $\eta^2_p = .15$. A Tukey HSD post hoc test indicated that participants both in the low history conditions ($M = 1.52$, $SD = 2.70$) and those in the high history/high target volunteer knowledge conditions ($M = 0.22$, $SD = 0.64$) volunteered significantly less time than those in the high history/low target volunteer knowledge conditions ($M = 4.47$, $SD = 6.89$; $p = .02$, Cohen's $d = 0.56$, and $p = .001$, Cohen's $d = 0.87$, respectively). No significant differences between the low history and the high history/high target volunteer knowledge conditions were found ($p = .46$). The two-way ANOVA revealed no other significant effects (ps > .1).

The one-way ANOVA showed significant differences between conditions, $F(5, 96) = 3.89$, $p = .003$, $\eta^2_p = .17$. Further contrasts showed that participants in the low harm/high history/high target volunteer knowledge condition (deterrence priming) volunteered significantly less time ($M = 0$, $SD = 0$) than those in both the low harm/high history/low target volunteer knowledge ($M = 5.72$, $SD = 9.04$), $t(17.00) = -2.69$, $p = .02$, Cohen's $d = 0.89$, and the low/low conditions ($M = 0.94$, $SD = 1.6$), $t(16.00) = -2.43$, $p = .03$, Cohen's $d = 0.83$, suggesting our deterrence hypothesis. On the other hand, participants in the low harm/high history/low target volunteer knowledge condition volunteered significantly more hours than those in the low harm/high history/high target volunteer knowledge and those in the low/low condition, $t(18.12) = -2.21$, $p = .04$, Cohen's $d = 0.82$. Our incapacitation hypothesis was thus disconfirmed. Finally, replicating the result of Study 1, participants in the high harm/low history condition (just deserts priming condition; $M = 2.11$, $SD = 3.35$) did not volunteer less time than those in the low/low ($p = .19$).

It is noteworthy that participants in the low harm/high history/low target volunteer knowledge condition volunteered substantial and significantly more time than those in all the other conditions, $F(1, 96) = 14.44$, $p < .001$, $\eta^2_p = .13$. Participants in the high harm/high history/low target volunteer knowledge ($M = 3.21$, $SD = 3.69$) also volunteered significantly more time than participants in both the high harm/high history/high target volunteer knowledge ($M = 0.44$, $SD = 0.81$), $t(20.06) = 3.19$, $p = .005$, Cohen's $d = 1.04$, and those in the low/low conditions, $t(25.11) = -2.48$, $p = .02$, Cohen's $d = 0.80$ (Figure 5.3.3.).

We created a dichotomized variable of the hours volunteered (1 = volunteered time and 0 = volunteered no time) to compare the proportions of participants who volunteered and did not volunteer time. No significant differences were found, ps > .1.

Figure 5.3.3.

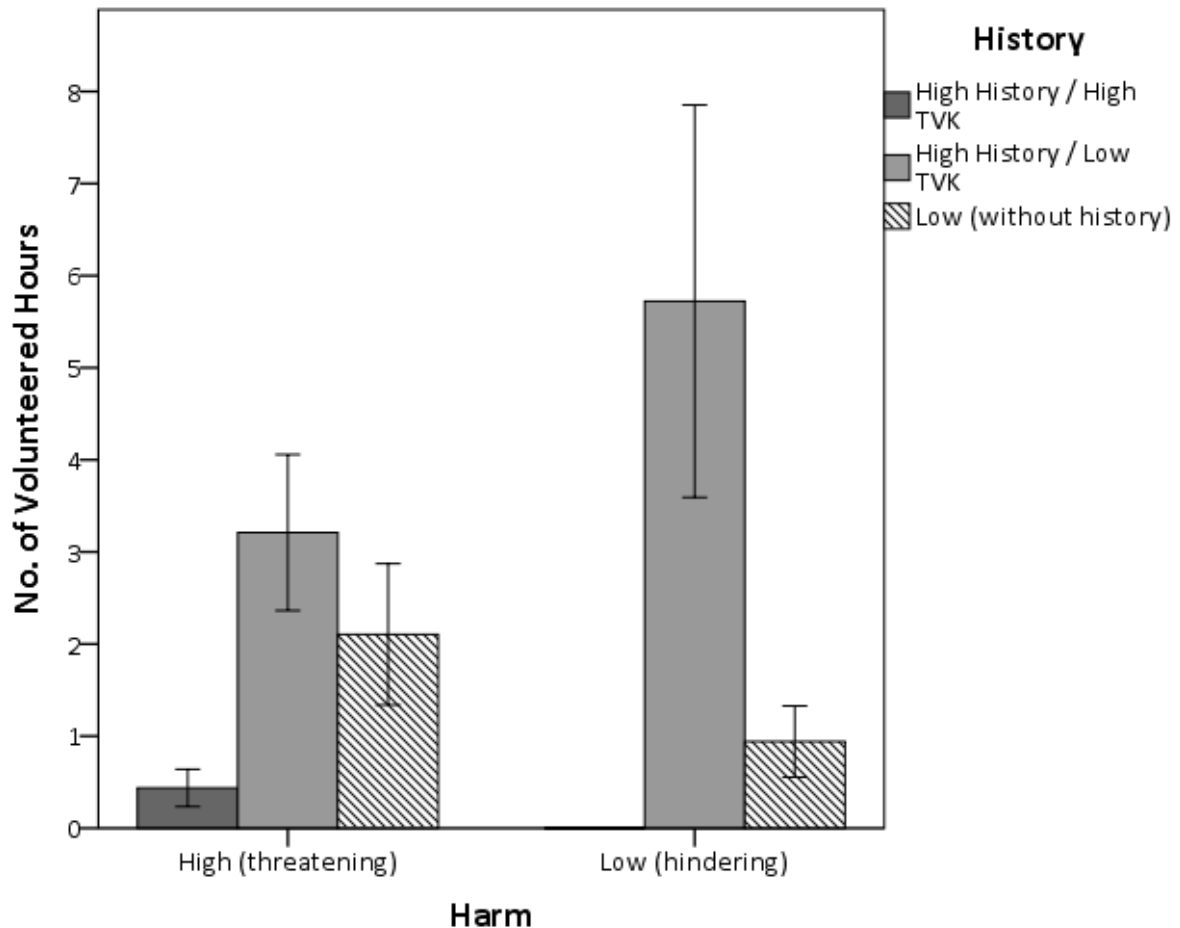*Number of Volunteered Hours as a Function of Harm, History, and Target Volunteer Knowledge (Study 2)*

Note. TVK: Target Volunteer Knowledge. Error bars represent standard errors.

Intentionality, responsibility, emotional stability, and compassion

Intentionality was subjected to the described two-way ANOVA, which showed a reliable main effect of the history, $F(2, 116) = 13.43$, $p < .001$, $\eta^2_p = .19$. A Tukey HSD post hoc test indicated that participants in the low history conditions considered the aggressions significantly less intentional ($M = 3.75$, $SD = 1.45$) than those of both the high history/high target volunteer knowledge ($M = 5.34$, $SD = 1.81$), $p < .001$, Cohen's $d = 0.97$, and those in the high history/low target volunteer knowledge conditions ($M = 5.39$, $SD = 1.56$), $p < .001$, Cohen's $d = 1.09$. No significant differences were found between the high history conditions ($p = .99$). The two-way ANOVA revealed no other significant effects ($ps > .10$). The one-way ANOVA showed significant differences between conditions, $F(5, 116) = 6.06$, $p < .001$, $\eta^2_p = .21$.

70

Further contrasts showed that participants in the high harm/low history (just deserts priming) condition rated the threatening behavior as less intentional ($M = 4.15$, $SD = 1.31$) than did those in all the other aggression conditions, $F(1, 116) = 9.07$, $p = .003$, $\eta^2_p = .07$, but did not differ from those in the low/low condition ($M = 3.35$, $SD = 1.49$).

Regarding the responsibility for the act, the two-way ANOVA resulted, like in the previous analyses, in a significant main effect of the history, $F(2, 114) = 4.84$, $p < .01$, $\eta^2_p = .08$. A Tukey HSD post hoc test showed that participants in the low history conditions considered the target less responsible for the aggressions ($M = 4.93$, $SD = 1.67$) than those of both the high history/high target volunteer knowledge ($M = 5.85$, $SD = 1.29$), $p = .02$, Cohen's $d = 0.62$, and those in the high history/low target volunteer knowledge conditions ($M = 5.76$, $SD = 1.30$), $p = .03$, Cohen's $d = 0.55$. No significant differences were found between the high history conditions ($p = .96$). The two-way ANOVA revealed no other significant effects ($ps > .1$). The one-way ANOVA revealed no significant differences between conditions ($p = .07$). Contrasts showed that participants in the high harm/low history (just desert priming; $M = 5.00$, $SD = 1.59$) attributed significantly less responsibility for the aggression than did those in both the high harm/high history conditions (high target volunteer knowledge: $M = 6.00$, $SD = 1.34$ and low target volunteer knowledge: $M = 5.86$, $SD = 1.28$), and those in the low harm/high history/low target volunteer knowledge ($M = 5.65$, $SD = 1.4$) condition, $F(1, 114) = 4.40$, $p = .04$, $\eta^2_p = .04$. No significant differences were found between the high harm/low history (just deserts priming) condition and both the low harm/high history/high target volunteer knowledge ($M = 5.67$, $SD = 1.28$), $F(1, 114) = 2.02$, $p = .16$, $\eta^2_p = .02$, and the low/low ($M = 4.85$, $SD = 1.79$), $F(1, 114) = 0.11$, $p = .74$, $\eta^2_p = 0.00$. The measures of intentionality and responsibility were highly correlated, $r = .64$, $p < .001$.

We also subjected the emotional stability variable to the referred two-way ANOVA. Results showed a reliable main effect of the harm magnitude, $F(1, 114) = 9.59$, $p = .002$, $\eta^2_p = .08$, indicating that participants in the high harm conditions (threatening behavior) considered the target as being more emotional stable ($M = 3.26$, $SD = 1.34$) than those in the low harm conditions (hindering behavior; $M = 2.56$, $SD = 1.06$). No other significant effects were found ($ps > .40$). Results of the one-way ANOVA revealed significant differences between conditions, $F(5, 114) = 2.30$, $p = .05$, $\eta^2_p = .09$. Further contrasts showed that participants in the high harm/low past offense (just deserts priming; $M = 3.2$, $SD = 1.36$) rated the target as being more emotional stable than those in the low/low condition ($M = 2.37$, $SD = 1.01$), $F(1, 114) = 4.49$, $p = .04$, $\eta^2_p = .04$. No other significant differences were detected ($ps > .10$).

Reported compassion for Cátia was high in all conditions ($M = 5.40$, $SD = 1.24$), and both the two-way and the one-way ANOVAs detected no significant differences between conditions.

Liking the target, beliefs about money and aggression, and past volunteering experience

To check whether there were differences between conditions in how much participants liked Cátia, we repeated the previously described two-way and one-way ANOVAs. The results of both analyses showed no significant differences between conditions ($ps > .10$). However, planned comparisons showed that participants in the high harm/low history (just deserts priming; $M = 4.85$, $SD = 1.04$) reported liking Cátia significantly more than participants in both the high harm/high history conditions (high target volunteer knowledge: $M = 4$, $SD = 1.34$, $t(37.49) = -2.27$, $p = .03$, Cohen's $d = 0.71$, and low target volunteer knowledge: $M = 3.9$, $SD = 1.79$, $t(32.75) = -2.11$, $p = .04$, Cohen's $d = 0.65$). No other significant contrasts were found ($ps > .10$). This measure was not significantly correlated with the volunteered time to help ($p = .15$).

Responses to the item measuring how much participants agreed with the sentence "In general, the more money a person has, the more people that person contacts or interacts with" were subjected to the same ANOVAs. Both ANOVAs showed no significant differences between conditions ($ps > .80$). The overall mean was 3.48, and the standard deviation was 1.80. We also subjected responses to the item "The more people an aggressive person contacts or interacts with, the more aggression victims that person causes." to the ANOVAs. Results of the two-way ANOVA showed a significant main effect of history, $F(2, 115) = 2.99$, $p = .05$, $\eta^2_p = .05$. A Tukey HSD post hoc test indicated that participants in the high history/high target volunteer knowledge conditions agreed more with the sentence ($M = 4.75$, $SD = 1.80$) than participants in the low history conditions ($M = 3.88$, $SD = 1.58$), $p = .04$, Cohen's $d = 0.51$. No significant differences were found between the high history/low target volunteer knowledge conditions ($M = 4.35$, $SD = 1.81$) and the high history/high target volunteer knowledge conditions, nor between the low history conditions ($ps > .40$). The one-way ANOVA resulted in no significant differences between conditions ($p = .14$). Responses to these items were significantly correlated ($r = .26$, $p = .004$). Participants in the low harm/high history/low target volunteer knowledge expressed agreement with both sentences "In general, the more money a person has, the more people that person contacts or interacts with" ($M = 3.75$, $SD = 1.73$) and "The more people an aggressive person contacts or interacts with, the more aggression victims that person causes." ($M = 3.95$, $SD = 1.82$).

Because responses to the past volunteering experience item varied significantly (ranging from 1 to 720 past volunteered hours), we decided to convert them into Z-scores before conducting the analyses. After this procedure, we conducted the described two-way ANOVA, which resulted in a reliable interaction between harm magnitude and history, $F(2, 57) = 3.47$, $p = .04$, $\eta^2_p = .11$. An analysis of simple effects showed that participants in the low harm/high history/low target volunteer knowledge condition had volunteered more time ($M = 0.31$, $SD = 1.09$), before participating in the study, than participants in both the low harm/high history/high target volunteer knowledge ($M = -0.44$, $SD = 0.19$) and those in the low harm/low history conditions ($M = -0.35$, $SD = 0.27$), $F(2, 57) = 3.40$, $p =$

.04, η^2_p = .11. No significant differences between the low harm/high history/high target volunteer knowledge and the low history conditions were found (p = .77). No other significant effects were found (ps > .40). The one-way ANOVA resulted in no significant differences between conditions (p = .13). The only significant contrast was found between participants in the low harm/high history/low target volunteer knowledge condition (M = 0.31, SD = 1.09) and participants in the low harm/high history/high target volunteer knowledge condition (M = -0.48, SD = 0.19), t(10.59) = -2.24, p = .05, Cohen's d = 1.01. The remaining contrasts were not significant (ps > .10). This variable was not correlated with the number of hours volunteered to help Cátia, p = .26.

We repeated the two-way ANOVA with the volunteered hours as dependent variable, now entering the participants' past volunteering experience as a covariate. After controlling for the past volunteering experience, there was still a significant main effect of history, F(2, 46) = 3.35, p = .04, η^2_p = .13. The two-way ANOVA revealed no other significant effects (ps > .20). We also repeated the one-way ANOVA with the volunteered hours as dependent variable, entering the participants' past volunteering experience as a covariate. With this analysis, however, after controlling for the past volunteering experience, the differences between conditions in the volunteered hours became non-significant (ps > .90).

See Table 5.3. for means and standard deviations of all the dependent variables.

Table 5.3.

Means and Standard Deviations for Punishment Severity, Perceived Seriousness of Need, Volunteer Hours, Intentionality, Responsibility, Emotional Stability, Compassion, Liking Cátia, Beliefs about Money and Aggression, and Past Volunteering Experience by Condition (Study 2)

Condition	Dependent Measure										
	Punishment severity	Perceived seriousness of need	Volunteer hours	Intentiona-lity	Responsibi-lity	Stability	Compassion	Liking Cátia	Belief about money[1]	Belief about money[2]	Past voluntee-ring experience (z-scores)
Low Harm/Low History – single hindering (n = 21)	3.05 (1.02) [a]	6.00 (0.82) [c]	0.94 (1.60) [ce]	3.35 (1.49) [a]	4.85 (1.79) [b]	2.37 (1.01) [b]	5.19 (1.33) [a]	4.67 (0.84) [a]	3.29 (1.98) [a]	3.81 (1.89)	-0.34 (0.30)
Low Harm/High History/Low Target Volunteer Knowledge – repeated hindering, impossible deterrence (n = 20)	3.37 (1.34) [ab]	6.50 (0.61) [b]	5.72 (9.04) [b]	5.25 (1.74) [b]	5.65 (1.39) [a]	2.85 (1.18) [ba]	5.8 (1.15) [a]	4.45 (1.57) [ac]	3.75 (1.73) [a]	3.95 (1.82)	0.31 (1.09)
Low Harm/High History/High Target Volunteer Knowledge – repeated hindering, possible deterrence (n = 20)	3.60 (1.35) [ab]	6.25 (0.64) [abc]	0 (0) [a]	5.15 (2.1) [b]	5.67 (1.24) [b]	2.47 (0.96) [b]	5.56 (0.78) [a]	4.80 (1.32) [a]	3.74 (1.94) [a]	4.79 (1.27)	-0.44 (0.19)
High Harm/Low History – single threatening (n = 20)	3.60 (0.88) [ab]	6.42 (0.69) [b]	2.11 (3.35) [cd]	4.15 (1.31) [a]	5.00 (1.59) [b]	3.2 (1.36) [a]	5.56 (0.92) [a]	4.85 (1.04) [a]	3.50 (1.73) [a]	3.95 (1.23)	-0.19 (0-73)
High Harm/High History/Low Target Volunteer Knowledge – repeated threatening, impossible deterrence (n = 21)	4.00 (1.30) [b]	6.45 (0.67) [b]	3.21 (3.69) [d]	5.52 (1.4) [b]	5.86 (1.28) [a]	3.38 (1.36) [a]	4.9 (1.7) [a]	3.90 (1.76) [c]	3.50 (1.91) [a]	4.75 (1.74)	-0.29 (0.29)
High Harm/High History/High Target Volunteer Knowledge – repeated threatening, possible deterrence (n = 21)	3.67 (1.62) [ab]	6.10 (0.28) [a]	0.44 (0.81) [e]	5.52 (1.5) [b]	6.00 (1.34) [a]	3.19 (1.36) [a]	5.45 (1.19) [a]	4.00 (1.34) [c]	3.33 (1.71) [a]	4.71 (1.52)	0.20 (1.23)

74

Note. Means that share a superscript do not differ significantly at p = .05, using paired t-tests.

[1] "In general, the more money a person has, the more people that person contacts or interacts with"

[2] "The more people an aggressive person contacts or interacts with, the more aggression victims that person causes."

Discussion

In Study 2 we predicted that participants in the conditions where the target acted aggressively multiple times would volunteer less time than those in conditions where the target acted aggressively a single time (H2.2.) This prediction was not confirmed. Participants in a condition where the target acted aggressively multiple times in a non-serious way, and where punishing with a deterrence motive was not possible neither volunteered less time than those in a condition where the target acted identically, but punishing with a deterrence motive was possible, nor than participants in a condition with a one-time non-serious aggressive act (thus disconfirming hypothesis H2.3.); however, when the target acted aggressively multiple times in a non-serious way and punishing with a deterrence motive was possible, participants volunteered less time than those both in a condition where the target acted identically, but punishing with a deterrence motive was not possible, and than participants in a condition with a one-time non-serious aggressive act – confirming H4.1..

As expected, participants volunteered to help less when deterring was possible than when it was not. In addition, when deterrence was possible and the target hindered multiple times, participants volunteered significantly less time than when the target hindered a single time, with a large effect size, suggesting they withheld help with a deterrence motive (although participants in the high history conditions did not punish the target, in the punishment severity item, more severely than in the low conditions – disconfirming H2.1. – which questions the activation of a utilitarian motive).

Interestingly, when deterrence was not possible, participants volunteered much more not only than when it was possible, but also than when the target hindered a single time, which was not predicted. This may be because in the other conditions participants read the same letter from the professor responsible for the study asking for volunteer support, which did not include the extra sentence highlighting that a deterrence motive was not possible. Only in the high history/low target volunteer knowledge conditions did participants have an extra sentence in the letter stating that Cátia would not know how many hours individuals, or the collective volunteered for her fundraising campaign. However, it is possible that reading in the letter that the target will not know of the volunteered time provided an opportunity for particularly strong personal pride at helping someone.

Participants in the low harm/high history/low target volunteer knowledge expressed clear agreement with both sentences "In general, the more money a person has, the more people that person contacts or interacts with" and "The more people an aggressive person contacts or interacts with, the more aggression victims that person causes.", thus confirming H2.4. However, hypotheses H2.1., H2.2., and H2.3. were not confirmed, failing to support incapacitation motivated help withholding with those beliefs (H2 in this study).

We previously found no evidence, in the Pretest and Study 1, that participants in the just deserts priming conditions withheld help. We reasoned that this could be due to not considering the hitting

behavior intentional and imputable, so in Study 2 we used a threatening behavior. However, participants in the just deserts priming condition (high harm/low history) did not volunteer less time than those in the low harm/low history condition, failing to support our prediction (H1.2.) and replicating the results obtained in the Pretest and Study 1. Moreover, even though participants considered the threatening aggression more serious than the hindering, it was not more severely punished (disconfirming H1.1.).

It is possible that, because the aggression in both the high harm magnitude/low history (just deserts priming) and the low/low (control) conditions occurred only once, and given the target's tragic situation – which can mitigate participant's tendency to find fault with and assign punishment to a target (Finkel, et al., 1996) – participants in these conditions might have simply forgiven the target (which possibly triggered more empathy felt for her). Supporting this interpretation, participants in these two conditions considered both the intentionality (pre-requisite for punishing) and the responsibility for the aggression significantly lower than in all the high history conditions (regardless of possible deterrence).

Analyses of participants' volunteering history indicated that the pattern of results of this variable closely resembles the pattern of volunteered hours in the study, but these variables were not significantly correlated. Still, it could be the case that participants' responses on this measure are influenced by how much they volunteered in the task – they could remember or estimate the amount of hours they generally volunteer differently depending on how much they did in this case; or the other way around: participants' volunteering history may have affected how much they volunteered in the study. Results of the analysis of covariance confirm the relationship between the volunteering history and the time volunteered in the study.

The letter given to participants in the high target volunteer knowledge conditions (containing the help measure) did not include information conveying that Cátia would learn the number of hours volunteered individually or in total (as opposed to the letter given to participants in the low target volunteer knowledge, which included information that stated that Cátia would not learn). This limitation was solved in a third study.

We reasoned that participants could also be willing to deter the target when she acted aggressively toward her colleague only one time. This possibility was addressed in the third study. In this study we also investigated whether participants' emotional responses could affect their willingness to help. Negative affective state theory (Cialdini et al., 1987) and the empathy-altruism hypothesis (Batson, 2011) claim that people may help a potential beneficiary who needs help because they feel emotional discomfort (the former theory), or because they feel an other-oriented feeling, congruent with the affective state of the potential beneficiary (the latter). Although these different emotional states should precipitate different motivations, both theories would predict that

77

information about a potential beneficiary's aggressive behavior would attenuate or inhibit the activation of the affective state, reducing the likelihood of helping.

With the addition of two items, Study 3 further examined the role of an incapacitation motive in withholding help to confirm whether or not participants withhold help basing on the belief that helping an aggressive person collect money would make more aggression victims.

5.4. STUDY 3

Aims and Hypotheses

In Study 3 we investigated whether participants would be willing to deter the target when she acted aggressively toward her colleague one time. We also examined the possible effect of target information on participants' emotional responses. The just deserts hypothesis (H1) was not tested, and so we did not include the harm magnitude factor in this study design; in all conditions, the target hindered her colleague.

We predicted that participants would withhold help as a form of punishment with a deterrence motive (H4), if we confirmed the following custom hypotheses, tested with planned contrasts:

- o A main effect of the target's past offense history on the punishment severity, with participants in the high history conditions punishing the target more severely than in the low conditions, indicating the activation of a utilitarian motive for punishing on the punishment severity item (H2.1.);

- o A main effect of the target's past offense history on helping, with participants in the high history conditions helping less than those in the low history conditions (H2.2.);

We also expected

- o Participants in a high history/high target volunteer knowledge condition volunteer less time than those in the high history/low target volunteer knowledge and those in both no-aggression conditions (deterrence after multiple aggressions; H4.1),

And/or

- o Participants in a low history/high target volunteer knowledge condition volunteer less time than those in the low history/low target volunteer knowledge and those in both no-aggression conditions (deterrence after a one-time aggression; H4.2.).

These results would indicate help withheld with a deterrence motive. To ensure this conclusion we added an item asking how much participants felt they had contributed to the target's learning regarding her behavior toward her colleagues. We also added a self-report measure of punishment motives, asking participants whether they thought the target deserved more help than the time they offered, and why or why not.

78

Finally, we expected that participants in all the aggression conditions would volunteer to help for less time than those in both no-aggression conditions, at least in part as result of lower empathic-concern and lower personal distress activation (H5).

Method

Participants

One hundred and nineteen undergraduate students participated in the study (aged 18-31; M = 20.14, SD = 2.44). Like in the previous studies, participants were asked to listen to and evaluate a bogus pilot radio program. They were randomly assigned to one condition of a 3 (history: no-aggression vs. low vs. high) X 2 (target volunteer knowledge: high vs. low) between-subjects design. Only female subjects participated. No participants were excluded from the sample due to suspicion about the veridicality of the recordings. We conducted power analysis using G*Power 3.1 (Faul, et al., 2009) to compute the power of a post-hoc ANOVA using a medium effect size (f = 0.25) and an error probability of 5% significance. The achieved power was of 51%.

Measures and procedure

The procedure was similar to previous studies. Participants listened to one of three versions of the radio program: one containing the description of the no-aggression condition, another in which the target threatened her colleague only one time, and a third condition in which she threatened multiple times. Deterrence was manipulated like in Study 2.

After listening to the recording, all participants were given a questionnaire (new in Study 3) in which they were asked to complete 23 emotional-reaction items (Batson, et al., 1997), indicating the intensity with which they felt each emotion toward the target (allowing measurement of empathy and negative-affective state). Then, in the same questionnaire, participants responded to the Affective dimension items of the State Empathy Scale (Shen, 2010). After that, participants were given different questionnaires, depending on whether they were in the no-aggression condition or in any of the aggression conditions. In the no-aggression condition they were given a second questionnaire with two items about irrelevant aspects of the scenarios (the same questionnaire given to participants in the no-aggression condition in Study 1), whereas participants in the aggression conditions responded to these items plus the punishment severity item (i.e., the questionnaire used in all the aggression conditions of the previous studies). In both cases, when the participant finished responding, the experimenter pretended to be looking for a third questionnaire and approached the participant with the envelope containing the letter with the measure of willingness to help and the manipulation of target volunteer knowledge, following the procedure used in previous studies. Unlike Study 2, the

letter allowing deterrence punishment (given to participants in high target volunteer knowledge conditions) included a sentence saying that Cátia would learn the number of hours volunteered individually or in total. The sentence was identical to the sentence included in letter given in the low target volunteer knowledge conditions, except that did not include the word "não" ("not") before the expression "terá conhecimento" ("will learn"). After reading the letter, participants were delivered a third questionnaire, whose items varied, again, depending on condition (no-aggression versus containing aggression information). Both were identical to the second questionnaires given in Study 1 for the no-aggression and the aggression information conditions. In the no-aggression condition this questionnaire included items about irrelevant aspects of the scenario, and the measures of Cátia's emotional stability, seriousness of her need situation, and compassion felt for her. Besides these questions, questionnaires delivered in the aggression information conditions also included measures of the intentionality and responsibility regarding Cátia's aggression, and manipulation checks. Once finished with the third questionnaire, participants received a fourth questionnaire. In all conditions, it included: a) the previously described measure of how much participants liked Cátia; b) an item asking participants how much they agreed with the sentence "The more money an aggressive person has, the more aggressive she might be toward people whom she contacts or lives with, thus possibly causing more aggression victims", to respond on a scale from 1 (*totally disagree*) to 7 (*totally agree*); c) an item asking participants how aggressive they believed Cátia was, on a scale from 1 (*not aggressive at all*) to 7 (*totally aggressive*); d) a self-report measure of the punishment motives, asking participants whether they thought the target deserved more help than the time they offered, and why or why not, giving some space in the questionnaire to write their response[5]; (e) an item asking what academic program (or major) participants were in. In all aggression conditions, we also had an item asking how much participants felt they had contributed to the target's learning regarding her behavior toward her colleagues, placed after the item asking how aggressive they considered the target. The scale for answering ranged from 1 (*I feel that I did not contribute at all*) to 7 (*I feel that I contributed totally*). Finally, in a final questionnaire participants completed the Cognitive and Associative dimension items of the State Empathy Scale (Shen, 2010), that were not included earlier to prevent influence on responses to other items. We included this scale with the purpose of exploring possible relations with skin conductance (SCL) data, but since we did not analyze SCL data, responses to the State Empathy Scale were not analyzed. For a summary of the measures used across all studies, see Table 5.4.1..

---

[5] We did not include a similar measure in our previous studies because it is known that people are sometimes unaware of the actual reasons that drive their decisions (Nisbett & Wilson, 1977), and we preferred to adopt a methodological paradigm that was already validated for the initial studies.

Table 5.4.1.

Summary of Measures given to Participants Across Studies

Studies	Conditions	Summary of measures given to participants (chronologically, from left to right)					
		1st questionnaire	Letter	2nd questionnaire	3rd questionnaire	4th questionnaire	5th questionnaire
Pre-test	All conditions	• Irrelevant aspects of the scenarios • Punishment severity	Measure of willingness to help	• Irrelevant aspects of scenarios • Seriousness of need • Harm magnitude and past offense manipulation checks			
Study 1	No-aggression condition	• Irrelevant aspects of the scenarios	Measure of willingness to help	• Irrelevant aspects of scenarios • Emotional stability • Seriousness of need • Compassion			
	Aggression information conditions	• Irrelevant aspects of the scenarios	Measure of willingness to help	• Irrelevant aspects of scenarios • Emotional stability			

		• Punishment severity		• Seriousness of need • Intentionality • Responsibility • Harm magnitude and past offense manipulation checks • Compassion	
Study 2	All conditions	• Irrelevant aspects of the scenarios • Punishment severity	Measure of willingness to help	• Irrelevant aspects of scenarios • Emotional stability • Seriousness of need • Intentionality • Responsibility • Harm magnitude and past offense manipulation checks • Compassion	• Liking for Cátia • Beliefs about money • Previous volunteering & hours
Study 3		1st questionnaire	2nd questionnaire	Letter	3rd questionnaire	4th questionnaire	5th questionnaire

No-aggression condition	- Emotional-reaction items - Affective dimension of State Empathy Scale	Irrelevant aspects of the scenarios	Measure of willingness to help	- Irrelevant aspects of scenarios - Emotional stability - Seriousness of need - Compassion	- Liking for Cátia - Beliefs about money - Cátia's aggressiveness - Deservingness of more help than offered - Previous volunteering & hours - Program attended	Cognitive and Associative dimensions of the State Empathy Scale
Aggression information conditions	- Emotional-reaction items - Affective dimension of State Empathy Scale	- Irrelevant aspects of the scenarios - Punishment severity	Measure of willingness to help	- Irrelevant aspects of scenarios - Emotional stability - Seriousness of need - Intentionality - Responsibility	- Liking for Cátia - Beliefs about money - Cátia's aggressiveness - Contribution to Cátia's learning - Deservingness of more help than offered	Cognitive and Associative dimensions of the State Empathy Scale

- Harm
  magnitude and
  past offense
  manipulation
  checks
- Compassion

- Previous
  volunteering &
  hours
- Program
  attended

Note. Sessions of all the studies started with participants signing the informed consent, attaching the SCL electrodes, and listening to the audio recording, and ended with the experimenter asking participants to place the letter in a box, stopping the SCL recording and removing the electrodes, probing for suspicion and providing debriefing (see details in the Method section of the Pre-test).

Manipulation checks

We subjected the likelihood of behaving improperly in the future variable to a 2 (history: high vs. low) X 2 (target volunteer knowledge: high vs. low) ANOVA. Results showed a significant main effect of the history, $F(1, 75) = 14.60$, $p < .001$, $\eta^2_p = .16.$, that participants in the high history conditions considered the target to be more likely to recidivate ($M = 5.17$, $SD = 1.60$) than those in the low history conditions ($M = 3.72$, $SD = 1.73$). No other significant results were found ($ps > .7$).

Hypothesis testing

Each following dependent variable was submitted to the same two-way ANOVA with 3 (history: no-aggression vs. low vs. high) X 2 (target volunteer knowledge: high vs. low). Results are described below.

Punishment motives

We predicted that we would obtain, on punishment severity, a main effect of history (H2.1.). The punishment severity item was submitted to the described ANOVAs. Because participants in the no-aggression conditions did not receive this item in their questionnaire, they were not included in these analyses. Results of the two-way ANOVA showed a main effect of history, $F(1, 72) = 7.1$, $p = .01$, $\eta^2_p = .09$, indicating that participants in the high history conditions punished the target more severely ($M = 4.03$; $SD = 1.3$) than those in the low history conditions ($M = 3.26$; $SD = 1.37$; Figure 9), confirming H2.1.. No other significant results were found ($ps > .10$). Planned comparisons showed that participants in the high history/low target volunteer knowledge condition assigned a more severe punishment to Cátia ($M = 4.44$, $SD = 0.51$) than participants in all the remaining conditions, $t(72) = -2.84$, $p = .006$, Cohen's $d = 0.94$ (Figure 5.4.1.).

Figure 5.4.1.

*Punishment Severity as a Function of History and Target Volunteer Knowledge (Study 3)*
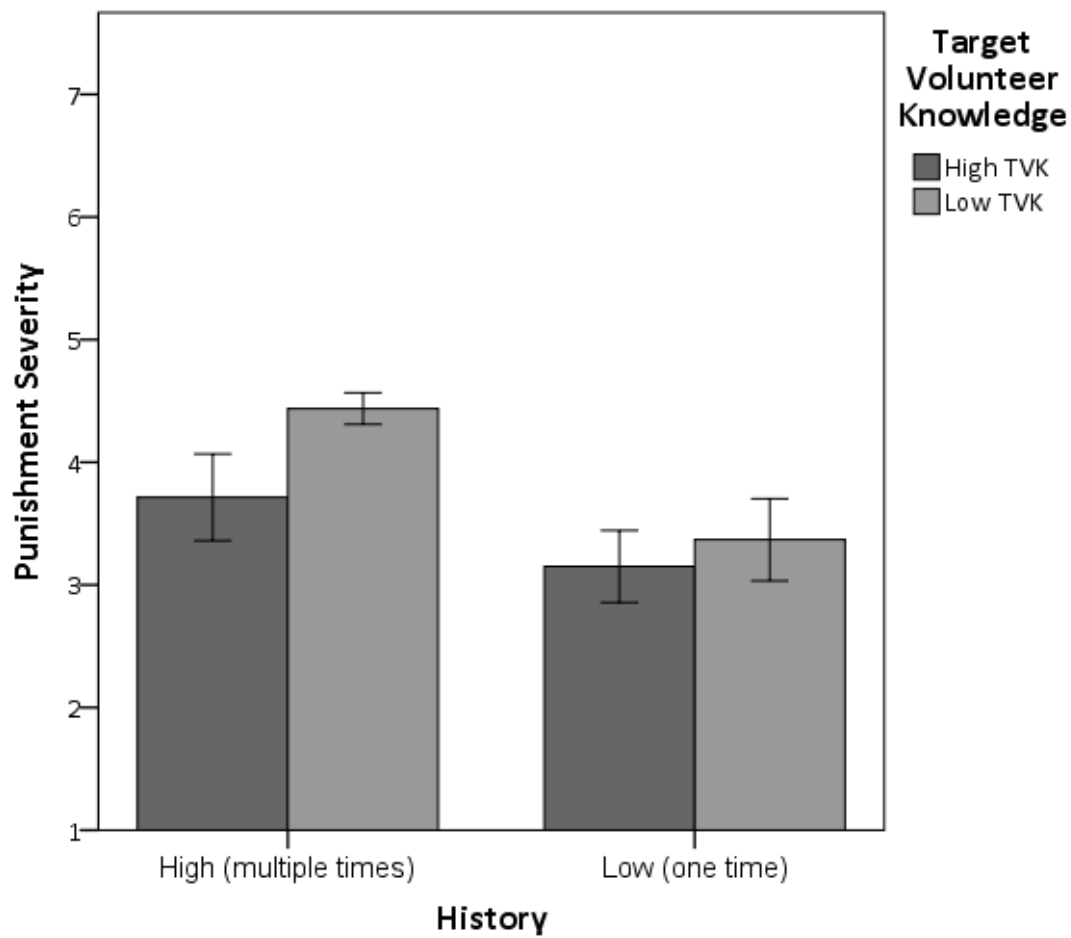
Note. TVK: Target Volunteer Knowledge. Error bars represent standard errors.

Perceived seriousness of need situation

The perceived seriousness of Cátia's need was high in all conditions (M = 6.34, SD = 0.67), and the two-way ANOVA showed no significant differences between conditions (Fs < 3; ps > .10) (Figure 5.4.2.).

Figure 5.4.2.

Perceived Seriousness of the Target's Need Situation as a Function of History and Target Volunteer Knowledge (Study 3)
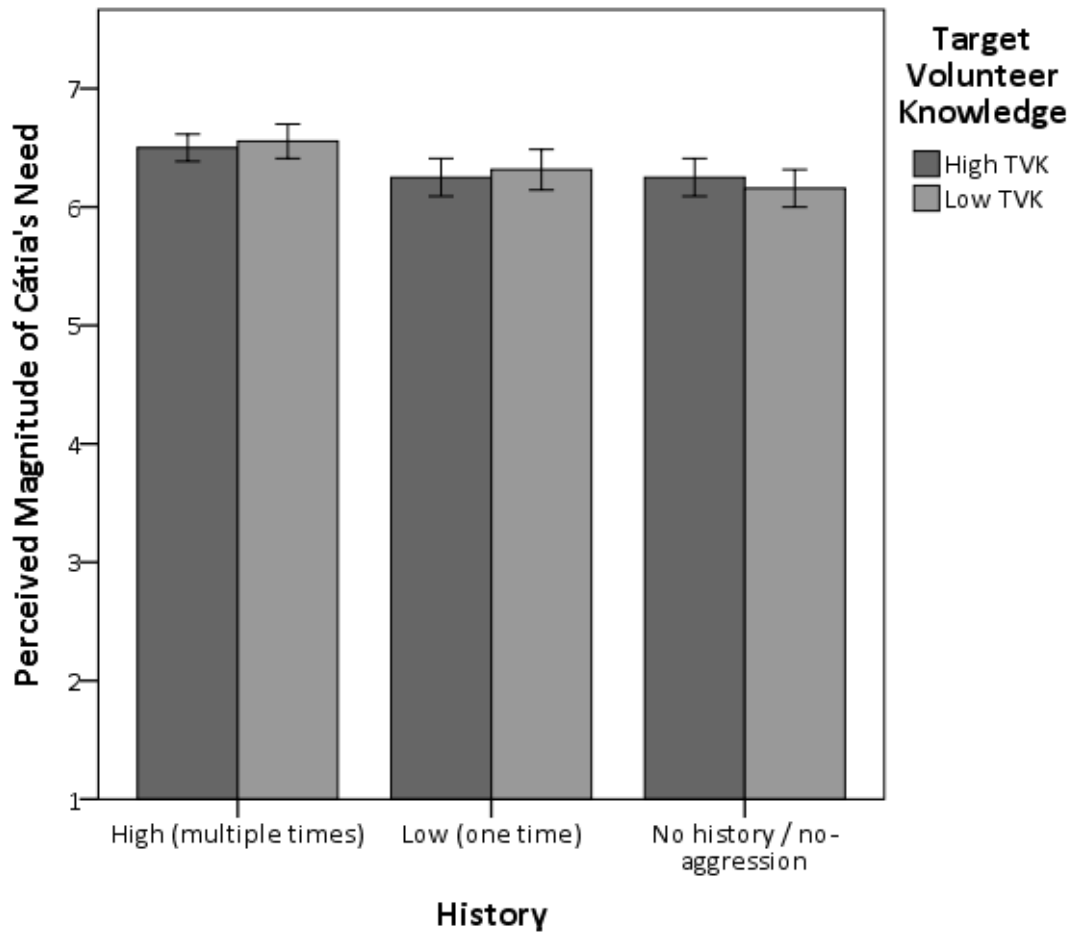
Note. TVK: Target Volunteer Knowledge. Error bars represent standard errors.

Volunteered hours (helping measure)

In this study we predicted a main effect of the target's past offense history on helping, with participants in the high history conditions helping less than those in the low history conditions (H2.2.). We also predicted that participants in a high history/high target volunteer knowledge condition would volunteer less time than those in the high history/low target volunteer knowledge and those in either no-aggression condition (deterrence after multiple aggressions; H4.1); alternatively, we predicted that participants in a low history/high target volunteer knowledge condition would volunteer less time than those in the low history/low target volunteer knowledge and those in either no-aggression condition (deterrence after a one-time aggression; H4.2.). Additionally, we also expected that participants in all the aggression conditions would volunteer to help less time than those in both no-aggression

conditions, at least in part as result of lower empathic-concern and lower personal distress activation (H5).

To test these hypotheses, we started by subjecting the volunteered hours to the described two-way ANOVA, which resulted in a reliable interaction between the target volunteer knowledge and history, $F(2, 95) = 4.35$, $p = .02$, $\eta^2_p = .08$. An analysis of simple effects showed that participants in the low history/low target volunteer knowledge condition volunteered less time ($M = 0.67$, $SD = 1.23$) than participants in the low history/high target volunteer knowledge condition ($M = 4.11$, $SD = 7.00$), $F(1, 95) = 8.77$, $p = .004$, $\eta^2_p = .08$. No significant difference in the volunteered time between the high history conditions was found ($p = .50$). Participants in the low history/high target volunteer knowledge also volunteered more time than participants in both the no-aggression/high target volunteer knowledge condition ($M = 1.26$, $SD = 1.86$), and those in the high history/high target volunteer knowledge condition ($M = 0.06$, $SD = 0.25$), $F(2, 95) = 7.04$, $p = .001$, $\eta^2_p = .13$. However, no significant difference in the volunteered time between the no-aggression/high target volunteer knowledge condition and the high history/high target volunteered knowledge conditions was found ($p = .39$). Moreover, no other significant results were found (ps > .10).

Planned comparisons indicated that participants in the high history/high target volunteer knowledge condition volunteered significantly less time ($M = 0.06$, $SD = 0.25$) than those in the high history/low target volunteer knowledge condition ($M = 0.82$, $SD = 1.43$; $t(17.05) = -2.17$, $p = .04$, Cohen's $d = 0.74$), and also volunteered less than those in both the no-aggression conditions: high target volunteer knowledge ($M = 1.26$, $SD = 1.86$; $t(18.78) = -2.78$, $p = .01$, Cohen's $d = 0.90$), and low target volunteer knowledge ($M = 1.6$, $SD = 2.39$; $t(14.29) = -2.49$, $p = .03$, Cohen's $d = 0.91$) conditions. Since the high history/high target volunteer knowledge condition corresponds best, from the conditions we created, to the deterrence sentencing, these results suggest that when it was possible to deter, participants may have activated a deterrence motive when volunteering time to help the target who had repeatedly hindered someone else's work. No other significant differences were found, ps > .10 (Figure 5.4.3.).

We used the dichotomized variable of the hours volunteered (1 = volunteered time and 0 = volunteered no time) to compare the proportion of participants who volunteered and did not volunteer time. We found that, only in the possible deterrence conditions (high target volunteer knowledge), the proportion of participants who withheld help (by volunteering no time at all) was significantly larger than the proportion of participants who helped by volunteering time, $\chi^2(2, n = 53) = 6.28$, $p = .04$. No other effects were significant.

Figure 5.4.3.

Number of Volunteered Hours as a Function of Aggression History and Target Volunteer Knowledge
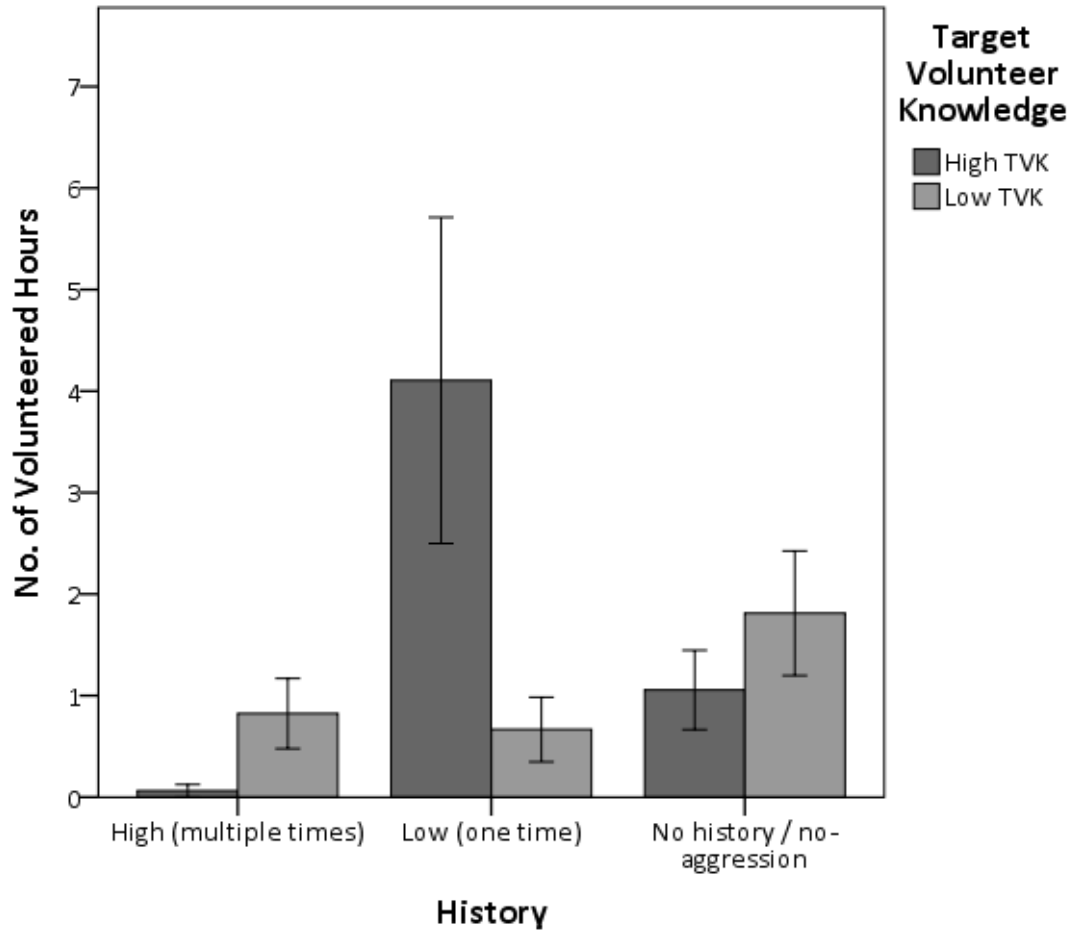
(Study 3)

Note. TVK: Target Volunteer Knowledge. Error bars represent standard errors.

Intentionality, responsibility, emotional stability, and compassion

The intentionality of the aggressive act by the target was subjected to the described two-way ANOVA, which showed a significant main effect of the history, $F(1, 75) = 11.69$, $p = .001$, $\eta^2_p = .14$ indicating that, consistent with results of Study 2, participants in the low history conditions considered the aggressions significantly less intentional ($M = 4.08$, $SD = 1.81$) than did those in the high history conditions ($M = 5.38$, $SD = 1.50$). The two-way ANOVA revealed no other significant effects ($ps > .10$).

Regarding the responsibility for the act, the two-way ANOVA resulted in a significant main effect of the target volunteer knowledge, $F(1, 75) = 9.60$, $p = .003$, $\eta^2_p = .11$, indicating that participants in the high target volunteer knowledge conditions considered the target less responsible for the aggression

(M = 4.68, SD = 1.52) than those in the low target volunteer knowledge (M = 5.66, SD = 1.24). The two-way ANOVA revealed no other significant effects (ps > .10).

The measures of intentionality and responsibility were significantly and moderately correlated, r = .36, p = .001.

The perceived emotional stability of the target was low, overall (M = 3.72, SD = 1.54). Results of the two-way showed no significant differences between conditions (ps > .70). Similarly, reported compassion for Cátia was high in all conditions, (M = 5.54, SD = 0.89), and the ANOVA showed no significant differences between conditions (ps > .40).

Liking the target, beliefs about money and aggression, and past volunteering experience

To check whether there were differences between conditions in how much participants liked Cátia, we subjected the liking Cátia variable to the described ANOVA. Results of the two-way ANOVA revealed a reliable main effect of history, $F(2, 101)$ = 3.60, p = .03, η^2_p = .07. Not surprisingly, a Tukey HSD post hoc test showed that participants in the low history conditions liked Cátia more (M = 4.86, SD = 0.83) than those in the high history conditions (M = 4.21, SD = 1.15), p = 0.02, Cohen's d = 1.00. No significant differences were found between the no-aggression conditions (M = 4.62, SD = 1.04) and the remaining conditions (ps > .20). The two-way ANOVA revealed no other significant effects (ps > .40). This measure was not significantly correlated with the volunteered hours variable (p = .34).

Responses to the item measuring how much participants agreed with the sentence "The more money an aggressive person has, the more aggressive she might be toward people whom she contacts or lives with, thus possibly causing more aggression victims" were subjected to the same ANOVA, finding no significant differences between conditions (ps > .30). The overall mean was 3.72, and the standard deviation was 1.93.

We also subjected responses to the item asking participants how aggressive they believed Cátia was to the ANOVA. Results of the two-way ANOVA showed a significant main effect of history, $F(2, 110)$ = 28.06, p < .001, η^2_p = .34. A Tukey HSD post hoc test indicated, as expected, that participants in the high history conditions considered Cátia significantly more aggressive (M = 3.41, SD = 1.38) than participants in both the low history conditions (M = 2.31, SD = 1.31), p < .001, Cohen's d = 0.82, and those in the no-aggression conditions (M = 1.47, SD = 0.69), p < .001, Cohen's d = 1.78. The low history conditions also significantly differed from the no-aggression conditions, p = .05, Cohen's d = 0.80. No other significant results were found (ps > .06). See Figure 5.4.4.

Figure 5.4.4.

Perception of Cátia's Aggressiveness as a Function of History and Target Volunteer Knowledge
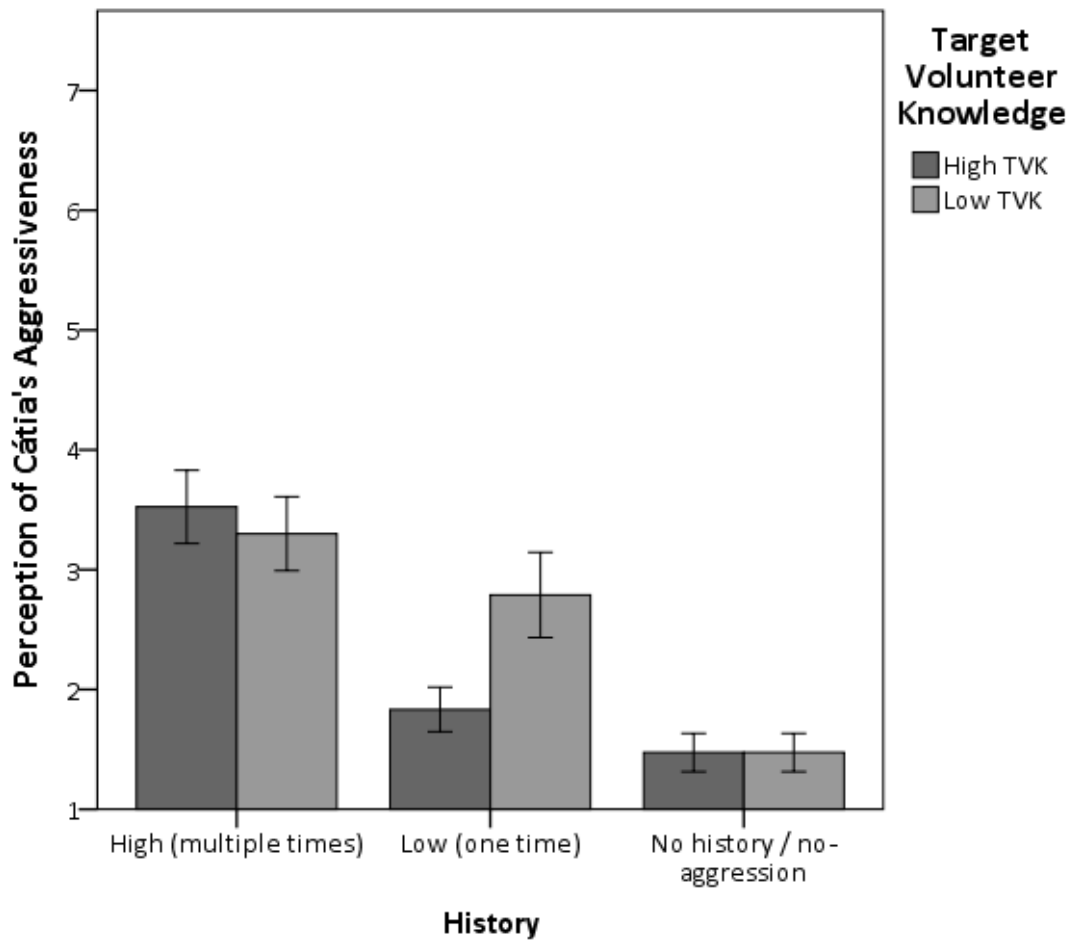
Note. TVK: Target Volunteer Knowledge. Error bars represent standard errors.

Regarding the past volunteering experience, no significant differences between conditions in participants' responses to the item asking whether they had volunteered in the past were found ($ps >$ .10). Most responses to the second item, which asked how many hours participants volunteered before the study, were not sufficiently objective, and we did not have a sufficient number of participants per condition to perform the ANOVA (less than 10 per condition). Neither measure was significantly correlated with the number of volunteered hours in the study (had volunteered: $r = .09$, $p = .37$; how many hours volunteered: $r = -.08$, $p = .63$).

Feeling of having contributed to Cátia's learning regarding her behavior

The reported feeling of having contributed to Cátia's learning regarding her behavior toward her colleagues was overall low (M = 2.78, SD = 1.74). The ANOVA showed no significant differences between conditions (ps > .50).

Self-report measure of punishment motives

Responses to the item that asked whether participants thought the target deserved more help than they offered, and why or why not, were coded in nine categories. Of these, eight categories represent responses saying that the target *did not deserve* more help than the participants volunteered, differing in terms of the justifications provided, and one category represent responses describing that the target *deserved* more help than they volunteered, for diverse irrelevant reasons (e.g., the target deserved more help, but they were not available for dispensing more), or providing no justification at all.

The eight categories of responses in which participants expressed that the target did not deserve more help than they gave varied according to the justification provided: (1) responses with justifications unrelated to the target's behavior (e.g., "because I think it is sufficient, she will already have many people helping her"); (2) justifications based on the target's behavior, but without clearly indicating that the time of help volunteered constituted a form of punishment for that behavior (there were no cases coded in this category); (3) justifications based on the target's behavior, and clearly indicating that the time of help volunteered was meant to be a form of punishment for the target's behavior, but without suggesting any particular punishment motive (e.g., "because she is not humble nor helps others"; "because she is vengeful"); for the following categories, there were no cases of responses: justifications based on the target's behavior, and clearly indicating that the help volunteered was meant to be a form of punishment for the target's behavior, specifically with (4) a retributionist motive, (5) a general utilitarian motive, (6) a deterrence motive, (7) an incapacitation motive, (8) both retributionist and utilitarian motives.

Most participants reported that the target deserved more help than they volunteered (84%). 13% of the participants considered that the target did not deserve more help than they offered due to irrelevant reasons, unrelated to the target's behavior, and only 3% of the participants reported that the target did not deserve more help than they offered, clearly indicating that such offered time was meant to be a form of punishment for the target's behavior, and without suggesting any particular punishment motive.

Emotional-reaction items

Responses to the 23 emotional-reaction items were subjected to a principal components analysis to understand the latent structure of the scale. The results revealed four components with eigenvalues

92

greater than 1.0, although a scree test indicated that a one-component solution was most appropriate (eigenvalue = 10.78), as found in previous research (Batson, et al., 1997). This first factor explained 47% of the total variance, with all the items loading 0.50 or higher, except *upset* (0.40) and *empathic* (0.40). An index of empathy was created by averaging the responses to all the emotional-reaction items (Cronbach's alpha = 0.94).

In this study we expected that empathic-concern and personal distress experienced by participants would mediate the relation between the target's aggressive behavior and participants' willingness to help. But because the components analysis resulted in items loading on a single empathic-concern factor (and not also on personal distress), a simple mediation analysis with empathic-concern as mediator was performed.

The indirect effect of the target's behavioral information on participants' volunteered time to help the target was not significant, b = -0.02, BCa CI [-0.09, 0.06] (Figure 5.4.5.).

Figure 5.4.5.

Mediation Model Testing the Effect of Target's Behavioral Information on Volunteered Time through Empathic Concern
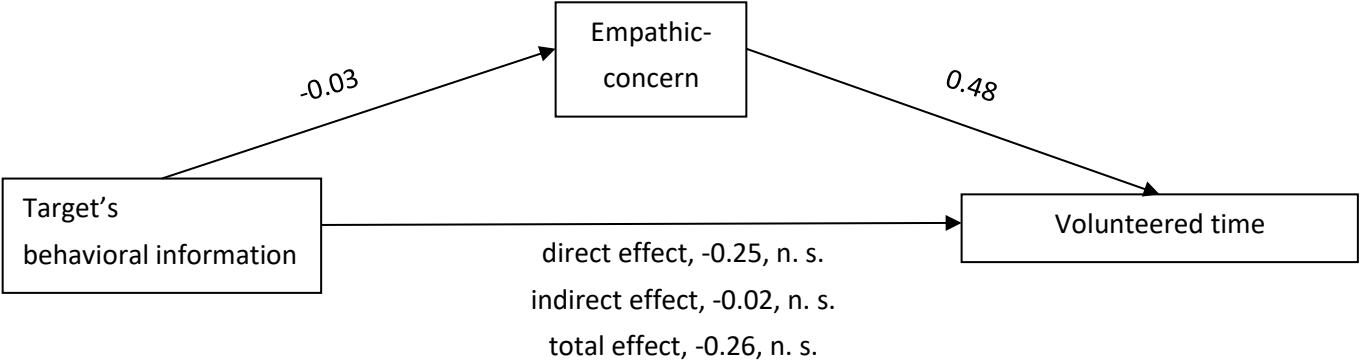
See Table 5.4.2. for means and standard deviations of all dependent variables, and Tables 7a, 7b, and 7c for the test results of the hypotheses of the studies.

Table 5.4.2.

Means and Standard Deviations for Punishment Severity, Perceived Seriousness of Need, Volunteered Hours, Intentionality, Responsibility, Emotional Stability, Compassion, Liking Cátia, Beliefs about Money and Aggression, Past Volunteering Experience, and Contribution to Cátia's Learning by Condition (Study 3)

Condition	Dependent Measure											
	Punishment severity	Perceived seriousness of need	Volunteer hours	Intentionality	Responsibility	Stability	Compassion	Liking Cátia	Belief about money[1]	Aggressiveness of Cátia[2]	Past volunteering experience	Contributed to Cátia's learning
No aggression/Low target volunteer knowledge (n = 18)	-	6.22 (0.65)[a]	1.60 (2.38)[acd]	-	-	2.37 (1.01)[a]	5.19 (1.33)[a]	4.89 (1.13)[ec]	4.39 (2.06)[a]	1.44 (0.71)[a]	63.20 (116.98)	-
No aggression/High target volunteer knowledge (n = 21)	-	6.19 (0.75)[a]	1.26 (1.86)[acd]	-	-	3.90 (1.51)[a]	5.70 (0.98)[a]	4.37 (0.90)[ea]	3.62 (1.88)[a]	1.50 (0.69)[a]	111.63 (160.52)	-
Low history/Low target volunteer knowledge (n = 19)	3.37 (1.46)[b]	6.32 (0.75)[a]	0.67 (1.23)[d]	4.05 (1.99)[a]	5.74 (1.33)[b]	3.84 (1.68)[a]	5.47 (0.80)[a]	4.71 (0.77)[edf]	3.63 (1.77)[a]	2.79 (1.55)[b]	89.33 (82.63)	2.83 (1.92)
Low history/High target volunteer knowledge (n = 20)	3.15 (1.31)[b]	6.25 (0.72)[a]	4.11 (7.00)[c]	4.10 (1.68)[a]	4.50 (1.50)[a]	3.80 (1.51)[a]	5.65 (0.79)[a]	5.00 (0.88)[cd]	3.11 (1.97)[a]	1.83 (0.79)[a]	57.86 (47.30)	2.89 (1.60)

High history/Low target volunteer knowledge (n = 20)	4.44 (0.51)[a]	6.56 (0.62)[a]	0.82 (1.43)[acd]	5.26 (1.52)[b]	5.58 (1.17)[bc]	3.47 (1.68)[a]	5.41 (1.12)[a]	4.18 (1.24)[af]	3.75 (2.25)[a]	3.30 (1.38)[b]	46.00 (65.69)	2.45 (1.47)
High history/High target volunteer knowledge (n = 21)	3.71 (1.62)[b]	6.50 (0.51)[a]	0.06 (0.25)[b]	5.48 (1.50)[b]	4.86 (1.56)[ac]	3.90 (1.76)[a]	5.35 (0.99)[a]	4.24 (1.09)[af]	3.86 (1.68)[a]	3.52 (1.40)[b]	63.00 (115.68)	2.95 (2.01)

Note. Means that share a superscript do not differ significantly at $p = .05$, using paired t-tests.

[1] "In general, the more money a person has, the more people that person contacts or interacts with"

[2] "The more people an aggressive person contacts or interacts with, the more aggression victims that person causes.

Table 5.4.3.

Results of Pretest and Study 1's Hypotheses Testing

	Help withheld with a just deserts motive (H1)			Help withheld with an incapacitation motive (H2)			Help withheld with both motives (H3)		
	Main effect of harm on punishment severity (H1.1.)	Main effect of harm on helping (H1.2.)	High harm/low history help less than no-aggression condition (H1.3.)	Main effect of history on punishment severity (H2.1.)	Main effect of history on helping (H2.2.)	Low harm/high history help less than no-aggression condition (H2.3.)	Main effects of both harm and history on punishment severity (H3.1.)	Main effects of both harm and history on helping (H3.2.)	High harm/high history help less than: high harm/low history; low harm/high history; no-aggression (H3.3.)
Pretest	No	No	No	No	No	No	No	No	No
Study 1	Yes	No	No	Yes	No	No (only low harm/high history)	Yes	No	No

Table 5.4.4.

Results of Study 2's Hypotheses Testing

	Help withhold with a just deserts motive (H1)		Help withhold with an incapacitation motive (H2)				Help withhold with a deterrence motive (H4)		
	Main effect of harm on punishment severity (H1.1.)	Main effect of harm on helping (H1.2.)	Main effect of history on punish severity (H2.1.)	Main effect of history on helping (H2.2.)	Low harm/high history/low TVK help less than low harm/high history/high TVK and low/low (H2.3.)	Agree with beliefs about money (H2.4.)	Main effect of history on punish severity (H2.1.)	Main effect of history on helping (H2.2.)	Low harm/high history/high TVK help less than low harm/high history/low TVK and low/low (H4.1)
Study 2	No	No	No	No	No	Yes	No	No	Yes

Note. TVK: Target Volunteer Knowledge.

97

Table 5.4.5.

Results of Study 3's hypotheses testing

	Help withhold with a deterrence motive (H4)				Withhold due to lower empathic-concern and personal distress (H5)
	Main effect of history on punish severity (H2.1.)	Main effect of history on helping (H2.2.)	Low history/high TVK help less than low history/low TVK and both no-aggression conditions (H4.1)	High history/high TVK help less than high history/low TVK and both no-aggression conditions (H4.2)	
Study 3	Yes	No	No	Yes	No

Note. TVK: Target Volunteer Knowledge.

Discussion

In Study 3 we predicted that participants in a low history/high target volunteer knowledge condition would volunteer less time than those in a low history/low target volunteer knowledge condition and those in both no aggression conditions (H4.2.). Results did not confirm this hypothesis, failing to find that participants withheld help with a deterrence motive when the target behaved aggressively one time. However, when the target had an history of aggressions (i.e., in the high history/high target volunteer knowledge condition) participants did volunteer less time than when the target behaved aggressively only one time (with a moderate effect size), and less than participants in both no aggression conditions (with large effect sizes) (H4.1.). Results obtained with our dichotomized variable (volunteered vs. did not volunteer) showed that the proportion of participants who withheld help was significantly larger than the proportion of participants who volunteered, but only when it was possible to deter (when the target would know how many hours were volunteered). Although we did not obtain a main effect of history on helping (H2.2.), these results are in line with the deterrence punishing literature, reiterating the importance of the perpetrator's history of offenses for deterring. Participants may have seen a single act of hindering as arising from situational factors rather than from the target's personality or habit that needs to be changed.

Curiously, results showed that when the target behaved aggressively one time, participants volunteered more hours to help her when it was possible to deter her, than when it was impossible. At the same time, there were no differences between the low target volunteer knowledge (impossible deterrence) conditions and the no aggression conditions. This pattern of results leaves us thinking that, instead of punishing, maybe participants were more motivated to help the target when they knew she would know who had helped – through the contact they left in the letter – and how much they offered to help.

We found no evidence that participants felt that they had contributed to the target's learning regarding her behavior toward her colleagues. As mentioned, our results suggest that participants withheld help with a deterrence motive. Accordingly, we expected that these participants would report having contributed to the target's learning in our measure – but it was not the case. One possible explanation is that participants were not aware that their help withheld could have had the underlying motive of deterring the target, i.e., they may have activated the motive unconsciously.

The fact that most participants reported that the target deserved more help than they offered suggests that this item may have triggered social desirable responses. Social desirability is the process by which people's need of approval can lead to responses aimed at satisfying that need (social desirability effect; Crowne & Marlowe, 1964). Specifically, in the case of the item, approval may have been sought by means of displaying a compassionate heart, stating that the target needs all possible help.

99

A similar mechanism may have intervened in the assignment of responsibility for the target's aggression by participants in the high target volunteer knowledge conditions: they considered it lower than participants in the low target volunteer knowledge conditions did possibly because they expected her to learn about their responsibility assignment (although they read that their responses were anonymous), so to obtain her approval.

Our hypothesis that the empathy felt by participants for the target would mediate the relation between the target's behavioral information and participants' willingness to help was not supported (H5). This result may be due to the also non-significant regressions obtained on our empathy index, and of the volunteered time, on the target's behavioral information. The significance of these regressions is required to establish mediation (Baron & Kenny, 1986). Thus, we found no evidence for the role of empathy in participants' decision to volunteer time to help the target.

Chapter 6. General Discussion

In this thesis we analyzed how the decision to help a target person could be influenced by aggressive behaviors the target enacted toward a third person in the past. Across the empirical studies conducted, we tried to understand whether young adults would withhold help from an aggressive target, and possible motives driving that decision. We investigated the reasons for thinking that some people might deserve more help than others.

One particularly relevant possible reason for withholding help due to the potential beneficiary's behavioral information, given its latitude of application, is to do it as a form of punishment with criminal justice motives.

One pretest and three studies were conducted in order to empirically test different criminal justice motives as a reason for withholding help. Each study formulated and tested specific hypotheses, with the aim of better clarifying the decision processes for helping and withholding help. In every study participants were given the opportunity to help a needy target. In some cases, that target had behaved aggressively toward a third person in the past. The study methods used allowed us to observe the activation of punishment motives underlying helping decisions. Following the described rationale, all the studies postulated that participants would volunteer less help when the target behaved aggressively than when she did not. Moreover, we assumed that objective patterns of responses – like assigning more severe punishments and volunteering less time by participants who listened to scenarios with relevant information for specific punishment motives – would indicate the activation of punishment motives supporting those decisions.

We showed that our manipulations of harm magnitude and historical frequency worked: participants considered our high harm behaviors of the target to cause more harm than our low harm behaviors and considered the target more likely to repeat the negative behaviors in the future in the high historical frequency conditions than in the low historical frequency conditions across all studies. Results from our punishment severity measure also supported the activation of the expected punishment motives in Studies 1 and 3: in Study 1, participants in the high harm magnitude conditions (including the just deserts priming) assigned a more severe punishment than those in the low harm, and participants in the high history conditions (including the incapacitation priming) punished the target more severely than those in the low. Likewise, in Study 3, participants in the high history conditions (including both incapacitation and deterrence priming) punished the target more severely than those in the low history conditions.

In the first study we found that participants who learned the target had hindered the work of her colleagues (a low-harm aggression) multiple times (high history) volunteered the least time to help her (a difference found with a large effect size; Cohen, 1988). We hypothesized that this might have been

done with a possible deterrence motive – i.e., expecting the target to make the association between the lack of help she received and her hindering behaviors, hoping to dissuade her from harming in the future.

In a subsequent partial replication of that study, Study 2, we used a threatening behavior instead of hitting, for the high harm magnitude conditions, with the goal of having a negative behavior that would be viewed as more intentional and imputable (as intentionality could have explained the unpredicted result of high time volunteered in the high harm conditions of Study 1). We also created conditions in which participants learned that deterrence by withholding help was impossible, as the target would not learn about the amount of help volunteered. The relevant results of Study 1 were reproduced: participants withheld help most when the target hindered the work of her colleagues multiple times and deterring was possible (a difference also found with a large effect size), and helped most when deterring was not possible. The theory that participants withheld help with a deterrence motive was thus supported.

In Study 3 we examined if participants might want to deter the target after only a single harm was committed as well as adding variables to examine whether emotions could be interfering in the helping decision process: it could be the case that affective arousal would increase the likelihood of helping. Results did not support the prediction that participants withheld help after a single harm was committed; however, they did try to deter the target via help withheld when she had an history of offenses (differences also found with moderate and large effect sizes), confirming the importance of this factor for deterrence sentencing found in previous research. Regarding the possible role of emotions, it was not possible in the study to establish an effect of empathy or affective arousal in participants' decision to volunteer time to help the target.

Regarding the possible activation of a just deserts motive when withholding help, the studies we conducted to address such a process – Studies 1 and 2 – provided no support. Results failed to corroborate prior studies that used similar paradigms (e.g., Darley, et al., 2000) where participants seemed to have more of a just deserts motive, based on the punishment patterns. However, it seems likely that this is because of the extremity of the need situation of the target we presented in the scenarios – a student whose parents died in a car accident the week before the recording date, needing money for food to support herself and her siblings. Future research should test the just deserts hypotheses with a target in a less extreme situation. The results of our experimental studies allow us to take the following conclusions:

a) behavioral information, specifically of aggressive behaviors a target person performed toward another in the past, can influence the decision taken by young adults of whether or not to act prosocially – amount of time volunteered to help – toward the target, when in clear need, extending previous findings with children (Vaish, et al., 2010);

b) people help a needy person less if he/she has acted aggressively toward another than if not;

c) people may punish a person in need by withholding help, when they know that the person has a history of aggressions (more than one occurrence), regardless of the magnitude of harm of the aggressive behaviors;

d) such punishment may be motivated by a deterrence belief, i.e., the belief that the target will associate the lack of help with his/her previous aggressive behaviors, learning that if he/she recidivates, the same punishment (or another) may result, causing the inhibition of possible new aggressions;

e) no evidence was found to support that people may withhold help motivated by a just deserts rationale of restoring justice.

The Pre-test and Studies 2 and 3 revealed particular conditions in which participants volunteered significantly more time to help the target than in the remaining conditions of the studies. In the Pre-test such result was attributed to considering the target's aggression (hitting) less intentional, which allowed forgiving that act, and to a possible higher motivation to relieve her distress by helping. In Study 1, we were able to confirm the lower intentionality of hitting when compared to hindering, and there was no longer a condition with significantly higher helping.

The conditions in which participants volunteered significantly more time to help the target were different between Studies 2 and 3 – but only in the history factor (they volunteered more in the high history and in the low history conditions, respectively), not in the target volunteer knowledge factor: in both studies participants volunteered significantly more time to help when they knew that the target would not learn how much they helped (i.e., in the low target volunteer knowledge conditions). As suggested in the discussion of Study 2, reading that the target would not learn how much participants helped may have triggered a strong feeling of pride in helping.

It is also possible that these differences in the volunteered time resulted from participants' past volunteering experience. In Study 2 in particular, when participants' past volunteering experience was included as a covariate in the analysis, the significant differences in volunteering time to help the target across conditions disappeared.

The conducted studies were not exempt of limitations. Our research is based on the assumption that withholding help from an aggressive person may be motivated by punishment for the perpetrated negative behavior. This assumption is based on children's behavior toward individuals who have behaved antisocially, which can consist of helping them less than non-antisocial individuals, and on interactions occurring in public good games, in which non-cooperative behaviors are punished to increase cooperation (see Chapter 1. Introduction). Following these patterns, our research showed that participants withheld help when the target had an history of offenses and it was possible to deter her. Our methods allowed us to infer that these participants were motivated to deter the target when

withholding help. But it is possible that other mental processes, which we were not able to detect with our measures, namely of an affective nature, automatic, or unconscious, may have driven their behavior, competing with our explanation. For example, the empathy felt for the target may have been curtailed by knowing about the previous aggressive behavior(s) of the target. Participants' past volunteering experience may also have exerted influence in their decisions to help the target. We attempted to test for these possible explanations, and did not find consistent evidence supporting them. In the former case, however, this may have been because of the lack of differentiation between empathic feeling and personal distress found in the measures and the limited sample size for a mediation analysis.

Another questionable point in our research concerns the extent to which theories derived from criminal justice, though used by lay people in criminal cases, may be activated when withholding help. As mentioned earlier, this second assumption is consistent with the transfer of knowledge and problem-solving strategies that people carry from one kind of problem to another kind, but we cannot be sure whether it was the case in our studies.

The research also had some methodological limitations. First, the sample we used in our pretest to select the aggressive behaviors was relatively small (n = 34). Second, in all the studies, the achieved power of the analyses was lower than what would be ideal (assuming 80% as the ideal). Likewise, the sample size used for testing the mediation model in Study 3 was also small (n = 101).

Third, in all studies, we manipulated history by including more information indicating that the target had an history of offenses in the recordings of the high history conditions than the specific information we provided to indicate that the target did not have an history of offenses, in the low history recordings.

Fourth, some aspects may have undermined the believability of the data collection.

Although the "pilot radio program" paradigm we used consistently triggered helping behavior in listeners, and we probed for suspicion in the end of the sessions, it is possible that some participants questioned the veracity of the alleged program, and the experimenter was not able to detect their doubts in the conversation dedicated to that. Some participants may have considered, for example, unlikely that someone whose parents died in the last week in an accident would give an interview in the following week, due to the recent emotional distress resulting from that event. Other participants may have felt too inappropriate or insensitive, in a circumstance like that, that the interviewer would ask the target to talk about an unfortunate episode at work (even though justifying the question by suggesting that such episode could have contributed to the fragile situation of the target). Similarly, when participants responded the first questionnaire, it could have sounded insensitive to ask how severe a punishment should be given the target for her improper behavior (given the seriousness of her need situation). Participants could have also been suspicious of the experimenter's performance

104

in telling them that he realized he was out of the second questionnaire, needing to go to print more copies, while they could read a letter that was supposed to be given when the session finished. Participants who expressed suspicion were only found in the Pre-test and Study 1. Two participants in the no-aggression conditions expressed suspicion; one from a low harm/low history condition; three from the low harm/high history conditions; and four from the high harm magnitude/high history conditions.

Although the paradigm used to trigger help assumes that listeners will help the target, it is questionable the extent to which they may direct their help not to the target, but to her siblings. In other words, to some extent, participants might help to benefit or increase the welfare of the target's siblings, more than the target's.

Our studies only utilized female participants, because empathy has been shown to be higher when participants were the same gender as the target. The conclusions we drew from results are valid for this population, but not necessarily for the male gender. The conclusions may have low ecological validity as well, since studies were only conducted in the laboratory and with undergraduate students.

Although we had a measure of how much participants liked the target in Study 3, we did not have one that measured how much participants identified themselves with the target.

Finally, previous research has shown that the harm magnitude of an aggression is less relevant than the aggressor's history of offenses for punishing with utilitarian motives (like the incapacitation and deterrence motives; Darley, et al., 2000). On the other hand, a person's past offenses are less important than the seriousness of a present offense for punishing with a just deserts motive.

However, it is reasonable to think that a serious aggression – a murder, for example – committed one time should also awake other people's motivation to incapacitate (e.g., by imprisonment) or safely dissuade the perpetrator, in order to prevent another possible grave deed. Likewise, knowing that a relatively unharmful offense – like shoplifting – was preceded by several similar occurrences, committed by the same long time recidivist may trigger one's motivation to make him pay for all that he stole – i.e., to suffer a just deserts punishment.

Thus, though the conducted studies were based on a validated empirical paradigm, we cannot be sure that each of these elements – harm magnitude and history of offenses – did not influence the activation of the punishment motives for which they have not been shown to be relevant.

Our qualitative measure designed to access more direct feedback from participants, regarding their intentions to punish and deter, probably triggered social desirable responses. It would be interesting to obtain a more personal and sincere account of participants' reactions and decisions facing the situation we presented to them. This might be accomplished utilizing an interview approach, for example. Interviewees would likely need to be briefed in the beginning of the sessions to emphasize

the importance of their honesty, since they could try to convey an image of being helpful people, in their answers, without being aware of that (because helping is socially desirable).

It would also be interesting to examine in further studies other possible mechanisms of withholding help caused by a potential beneficiary's behavioral information. One of the mechanisms of interest would be whether behavioral information may undermine beliefs in self-benefits resulting from help (see Chapter 4. Withholding prosocial behavior). Probably some indirect measure would be needed to address which such benefits one would expect from helping a target in a given situation, since reports of prosocial behavior tend to be contaminated by social desirability. One possible approach would be to measure how much participants expected to receive some self-benefit (and/or avoid some aversive situation) by helping, when confronted with a target in need, before and after learning that the target acted antisocially or aggressively toward a third person in the past. One might expect that the intensity of those beliefs would decrease from the first to the second measure.

Likewise, other described processes of withholding help due to behavioral information could be investigated, extending the present line of research:

- One study could partially replicate the studies of Fehr and Gachter (2000; 2002) with public good games. Their proposed processes – in which behavioral information elicits negative emotions which in turn cause punishment – could be tested in more natural settings, as well as with other prosocial behaviors different from giving and removing pecuniary resources, like helping to accomplish a task or comforting a target.

- It would also be interesting to examine the extent to which the motive to help may compete with the motive to punish, and how exactly that interfering effect may happen in a decision process (that is, after perceiving a potential beneficiary's need and knowing about some past negative behavior). The research paradigm that could be used to address this question could be similar to the one described previously for assessing the undermining of beliefs in self-benefits: after facing a target in need with a negative behavioral record, and after being presented with an opportunity to help, whether participants believed they would receive some self-benefit from helping could be measured, and whether they hesitated for some reason (i.e., some punishment motive). The disclosure of the negative behavioral information could be delayed for a varying amount of time before or after the exposure to the state of need.

- Another mechanism that could be investigated is the trust in the potential benefactor's need. One may perceive the potential beneficiary's need, but without believing in such perception (see Chapter 4. Withholding prosocial behavior). This could be assessed by means of a self-report measure, e.g., an item in a questionnaire asking participants how much they believed in the need they perceived (after asking them how serious they considered a target's need, as used in the conducted studies).

106

- Measuring whether participants would withhold help because of devaluing self-benefits possibly resulting from helping would require accessing participants' motives to help and how important they were for them. This probably would not be achieved through direct measures, since participants would presumably tend to state altruistic motives important for them (due to social desirability). One possible solution could be, for example, have participants read, in a questionnaire, that most people tend to help for many reasons, with altruism, reportedly, being the most important; and also that, besides that reason, people tend to help for many other egoistic reasons. Then participants would be asked to rate the importance they attributed to each egoistic helping motive of a list, if they would help with any of such motives (as well as given the opportunity to add another egoistic motive, not listed). The idea of this approach would be to convey to participants the false idea that *everyone helps with egoistic motives*, and that *it is common*, thus lowering their guard, i.e., inhibiting their possible need of responding according to the social expectation of not appearing egoistic. Having experimental sessions run with a single participant at a time, each participant would then complete a short distracting task, while the experimenter would briefly analyze his/her responses and inform a confederate actor about the participant's most valued self-benefit. Once completed the distracting task, the participant would be asked to complete a third task with an ostensive co-participant, who would be the confederate. The confederate would then introduce him/herself, conveying information that would decrease the likelihood of helping him/her with the most valued self-benefit reported by the participant (e.g., if the participant indicated that they value most obtaining social approval, the confederate would introduce him/herself as a drug addict). In the middle of the task, the confederate would suddenly drop a box of pencils (or express another unexpected need), and if the participant withheld help, that would indicate that he/she withheld help due to devaluing his/her most valued self-benefit.

Another route that could elucidate help withhold could come from using physiological measures. Skin conductance covaries with empathy (Marci & Orr, 2006). Feeling less empathy for an aggressive target, as measured by changes in levels of skin conductance, could explain less willingness to help that target. The skin conductance data (SCL) collected in the studies described were not reliable due to not using a required filter in the recording (which we only understood after completing the data collection of Study 3), and so it was not possible to explore its role in helping. Future research into withholding help may benefit from using skin conductance data (and possibly from other physiological measures), particularly because it is an indirect measure, which eliminates the social desirability often occurring with direct measures of helping.

Chapter 7. Conclusion

The findings and conclusions of the research conducted in the scope of the present thesis reiterate the importance of punishment often discussed in past research (Carlsmith, 2006) by showing how it may be administered out of a concern for others, accounting for its regulatory function in social life. But most importantly, the findings suggest that punishment behavior might assume a more subtle, more socially acceptable – and certainly more self-protective in diverse aspects – form through help withheld.

The work in this thesis also contributes significantly to clarifying a seldom researched, but essential feature of prosocial behavior: deservingness and selectivity mechanisms. It helps explain hesitancy, delay, avoidance (many times reasoned), preference, and inaction – or alternative action – when one becomes aware of an opportunity to help, in any situation. Consider these examples:

- Your newcomer neighbor is moving in, and you are available to help, but you just watched him/her insulting the moving company's employee for accidently dropping a book.
- Your cousin refused to come visit your sick father in the hospital, with some excuse, and now asks you to help him in gardening next weekend.
- A colleague of yours who keeps sabotaging the work of another colleague because he wants to get an announced promotion – and now he kindly asks your advice for increasing his performance.

If you think you would feel hesitant in helping your neighbor picking up the book, it is possibly because your neighbor's aggressive behavior toward the employee may be one such that you would not want to see again (toward the actual or another potential victim); and you may believe that by withholding help toward your neighbor, he could associate such punishment to his previous condemnable demeanor, decreasing his willingness to repeat it[6]. According to our studies results, this would be a likely process for your hesitancy (particularly if you knew that the aggression in question was preceded by other analogous behaviors).

Likewise, if you think you would not feel available to help your cousin after his refusal to visit your father, our results suggest that you more probably would be motivated to change his behavior than to give him a just deserts. Still, if on the other hand you would operate more under a retributionist perspective and you considered the excuse provided by your cousin as valid (decreasing the intentionality of your cousin's refusal), that could influence your reaction, perhaps leading you to forgive him and help him gardening.

---

[6] (i.e., you would *believe* that your neighbor would anticipate another punishment when in need of help, by withholding help or other, should he repeat an aggression)

If you think, in the third example, that you would rather withhold your advice to your colleague, that could also be explained, following our results, by activating a deterrence motive, or, with a lower probability or extent, a just deserts motive. But in addition to that, you could also feel that your colleague should be impeded from continuing undermining your other colleague's work. You could also believe that, once promoted, your colleague could behave in similar ways toward other people in the company to get what he wanted. Thus, you could think that giving him advice to improve his performance would increase his chances to be promoted, which would in turn cause more victims of his conduct. If you withheld your advice having this in mind, you would be punishing your colleague with an incapacitation motive. However, our results do not suggest that this response would be likely.

Studies presented in the empirical section explore moderators and mediators that might explain differences in helping behavior.

More comprehensive knowledge of interpersonal relationships may be of particular importance in contexts where people have limited resources to help others and deservingness becomes a salient criterion for social support. Understanding when and why prosocial behavior is withheld may enable people and organizations to counteract people's tendency toward inaction. This work may increase the understanding of, for instance, the lack of involvement in social support institutions, such as those that work with delinquent youth, addicts, or the poor. This would provide information that could be used to make fundraising or awareness campaigns more effective. Such benefit is aligned with the UN's sustainable development goal of promoting inclusive societies, providing access to justice for all, and build effective, accountable and inclusive institutions at all levels (Goal 16; Department of Economic and Social Affairs of the United Nations, n. d.). Society's particular groups like the referred ones may be subjected to more or less overt punishment rationales of decision makers, which may inhibit, for example, their willingness to provide those groups equalitarian access to justice. Informing decision makers of how they may inadvertently activate punishment motives toward some individuals, and how it may interfere in their decisions, namely under the form of help withheld, can help them be fairer and more transparent in those decisions. In the same sense, such knowledge may also contribute to ensure the fulfilment of non-discriminatory laws and policies (which is the Target 16.b of the mentioned UN's Goal 16) as well as, consequently, to the protection of fundamental freedoms (Target 16.10 of UN's Goal 16). These goals are consonant with the European Union general aims of combating social exclusion and discrimination, and promoting social justice and protection (European Commission, Directorate-General for Communication, n. d.).

It may also assist organizations with conflict between employees by helping explain possible individual psychological responses to another's antisocial behavior.

This work sheds light on how helping decisions and behaviors may be conditional on whom we may help.

109

References

Ahern, J., Stuber, J., & Galea, S. (2007). Stigma, discrimination and the health of illicit drug users. *Drug and Alcohol Dependence, 88*(2-3), 188–196. https://doi.org/10.1016/j.drugalcdep.2006.10.014

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53*(1), 27–51. https://doi.org/10.1146/annurev.psych.53.100901.135231

Anderson, M. C., & MacCoun, R. J. (1999). Goal conflict in juror assessments of compensatory and punitive damages. *Law and Human Behavior, 23*, 313–330. https://doi.org/10.1023/A:1022308515445

Andreu, J., Ramirez, J., & Raine, A. (2006). Un modelo dicotómico de agresión y su evaluación mediante dos autoinformes: El CAMA y el RPQ. *Psicopatología Clínica, Legal y Forense, 5*, 25-42.

Azrin, N. H., & Holz, W. C. (1966). Punishment. In W. R. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 213-270). Appleton-Century-Crofts.

Axelrod, R. (1980). More effective choice in the Prisoner's Dilemma. *Journal of Conflict Resolution, 24*(3), 379–403. https://doi.org/10.1177/002200278002400301

Banks, C. (2017). *Criminal justice ethics: Theory and practice*. Sage Publications, Inc..

Baron, R. A. (1997). The sweet smell of…helping: Effects of pleasant ambient fragrance on prosocial behavior in shopping malls. *Personality and Social Psychology Bulletin, 23*(5), 498–503. https://doi.org/10.1177/0146167297235005

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*(7), 462–479. https://doi.org/10.1037/0003-066X.54.7.462

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Batson, C. D. (1994). Why act for the public good? Four answers. *Personality and Social Psychology Bulletin, 20*(5), 603–610. https://doi.org/10.1177/0146167294205016

Batson, C. D. (2011). *Altruism in humans*. Oxford University Press.

Batson, C. D., Ahmad, N., Lishner, D. A., & Tsang, J.-A. (2002). Empathy and altruism. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (p. 485–498). Oxford University Press.

Batson, C. D., Ahmad, N., Powell, A. A., & Stocks, E. L. (2008). Prosocial motivation. In J. Y. Shah & W. L. Gardner (Eds.), *Handbook of motivation science* (p. 135–149). The Guilford Press.

Batson, C. D., Ahmad, N., & Stocks, E. L. (2011). Four forms of prosocial motivation: Egoism, altruism, collectivism, and principlism. In D. Dunning (Ed.), *Social motivation* (pp. 103-126). Psychology Press.

Batson, C. D., Ahmad, N., Yin, J., Bedell, S. J., Johnson, J. W., Templin, C. M., & Whiteside, A. (1999). Two threats to the common good: Self-interested egoism and empathy and empathy-induced altruism. *Personality and Social Psychology Bulletin, 25*(1), 3–16. https://doi.org/10.1177/0146167299025001001

Batson, C. D., Batson, J. G., Griffitt, C. A., Barrientos, S., Brandt, J. R., Sprengelmeyer, P., & Bayly, M. J. (1989). Negative-state relief and the empathy—altruism hypothesis. *Journal of Personality and Social Psychology, 56*(6), 922–933. https://doi.org/10.1037/0022-3514.56.6.922

Batson, C. D., Batson, J. G., Slingsby, J. K., Harrell, K. L., Peekna, H. M., & Todd, R. M. (1991). Empathic joy and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology, 61*(3), 413–426. https://doi.org/10.1037/0022-3514.61.3.413

Batson, C. D., Batson, J. G., Todd, R. M., Brummett, B. H., Shaw, L. L., & Aldeguer, C. M. R. (1995). Empathy and the collective good: Caring for one of the others in a social dilemma. *Journal of Personality and Social Psychology, 68*(4), 619–631. https://doi.org/10.1037/0022-3514.68.4.619

Batson, C. D., Bolen, M. H., Cross, J. A., & Neuringer-Benefiel, H. E. (1986). Where is the altruism in the altruistic personality? *Journal of Personality and Social Psychology, 50*(1), 212–220. https://doi.org/10.1037/0022-3514.50.1.212

Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology, 40*(2), 290–302. https://doi.org/10.1037/0022-3514.40.2.290

Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imagining how you would feel. *Personality and Social Psychology Bulletin, 23*(7), 751–758. https://doi.org/10.1177/0146167297237008

Batson, C. D., Eklund, J. H., Chermok, V. L., Hoyt, J. L., & Ortiz, B. G. (2007). An additional antecedent of empathic concern: Valuing the welfare of the person in need. *Journal of Personality and Social Psychology, 93*(1), 65–74. https://doi.org/10.1037/0022-3514.93.1.65

Batson, C. D., & Powell, A. A. (2003). Altruism and prosocial behavior. In T. Millon, & M. J. Lerner (Eds.), *Handbook of psychology: Personality and social psychology* (pp. 463-484). John Wiley & Sons.

Bekkers, R., & Ottoni–Wilhelm, M. (2016). Principle of care and giving to help people in need. *European Journal of Personality, 30*(3), 240–257. https://doi.org/10.1002/per.2057

Bentham, J. (1973). *Principles of morals and legislation*. Anchor Books

Berkowitz, L. (1972). Social norms, feelings, and other factors affecting helping behavior and altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology*: Vol. 6 (pp. 63–108). Academic Press.

Berkowitz, L. (1973). Reactance and the unwillingness to help others. *Psychological Bulletin, 79*(5), 310–317. https://doi.org/10.1037/h0034443

Berkowitz, L., & Daniels, L. R. (1963). Responsibility and dependency. *The Journal of Abnormal and Social Psychology, 66*(5), 429–436. https://doi.org/10.1037/h0049250

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin, 88*(1), 1–45. https://doi.org/10.1037/0033-2909.88.1.1

Boster, F., Fediuk, T., & Kotowski, M. R. (2001). The effectiveness of an altruistic appeal in the presence and absence of favors. *Communication Monographs, 68*(4), 340-346. https://doi.org/10.1080/03637750128074

Brickman, P., Folger, R., Goode, E., and Schul, Y. (1981). Microjustice and macrojustice. In M. Lerner (Ed.), *The justice motive in social behavior* (pp. 173–204). Plenum.

Branch, M. N. (2000). Punishment. In A. E. Kazdin (Ed.), *Encyclopedia of psychology*. American Psychological Association.

Brown, J. D., & Smart, S. A. (1991). The self and social conduct: Linking self-representations to prosocial behavior. *Journal of Personality and Social Psychology, 60*(3), 368–375. https://doi.org/10.1037/0022-3514.60.3.368

Buckholtz, J. W., & Marois, R. (2012). The roots of modern justice: Cognitive and neural foundations of social norms and their enforcement. *Nature Neuroscience, 15*(5), 655–661. https://doi.org/10.1038/nn.3087

Carlo, G., Eisenberg, N., Troyer, D., Switzer, G., & Speer, A. L. (1991). The altruistic personality: In what contexts is it apparent? *Journal of Personality and Social Psychology, 61*(3), 450–458. https://doi.org/10.1037/0022-3514.61.3.450

Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology, 42*(4), 437-451. https://doi.org/10.1016/j.jesp.2005.06.007

Carlsmith, K. M., & Darley, J. M. (2008). Psychological aspects of retributive justice. In M. P. Zanna (Ed.), *Advances in experimental social psychology*: Vol. 40 (pp. 193–236). Elsevier Academic Press. https://doi.org/10.1016/S0065-2601(07)00004-4

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology, 83*(2), 284–299. https://doi.org/10.1037/0022-3514.83.2.284

Carpenter, J. P., & Matthews, P. H. (2004). *Social reciprocity* (No. 1347). IZA Discussion Papers.

Carpenter, J. P., Matthews, P. H., & Ong'ong'a, O. (2004). Why punish? Social reciprocity and the enforcement of prosocial norms. *Journal of Evolutionary Economics, 14*(4), 407–429. https://doi.org/10.1007/s00191-004-0212-1

Cavender, G. (1979). Special deterrence: An operant learning evaluation. *Law and Human Behavior, 3*(3), 203–215. https://doi.org/10.1007/BF01039791

Clary, E. G., Snyder, M., Ridge, R. D., Copeland, J., Stukas, A. A., Haugen, J., & Miene, P. (1998). Understanding and assessing the motivations of volunteers: A functional approach. *Journal of Personality and Social Psychology, 74*(6), 1516–1530. https://doi.org/10.1037/0022-3514.74.6.1516

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy–altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology, 73*(3), 481–494. https://doi.org/10.1037/0022-3514.73.3.481

Cialdini, R. B., Darby, B. L., & Vincent, J. E. (1973). Transgression and altruism: A case for hedonism. *Journal of Experimental Social Psychology, 9*(6), 502-516. https://doi.org/10.1016/0022-1031(73)90031-0

Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J., & Beaman, A. L. (1987). Empathy-based helping: Is it selflessly or selfishly motivated? *Journal of Personality and Social Psychology, 52*(4), 749–758. https://doi.org/10.1037/0022-3514.52.4.749

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.

Coke, J. S., Batson, C. D., & McDavis, K. (1978). Empathic mediation of helping: A two-stage model. *Journal of Personality and Social Psychology, 36*(7), 752–766. https://doi.org/10.1037/0022-3514.36.7.752

Cramer, R. E., McMaster, M. R., Bartell, P. A., & Dragna, M. (1988). Subject competence and minimization of the bystander effect. *Journal of Applied Social Psychology, 18*(13), 1133–1148. https://doi.org/10.1111/j.1559-1816.1988.tb01198.x

Crocker, J. & Bylsma, W. H. (1996). Self-esteem. In A. S. R. Manstead and M. Hewstone (Eds.), *Blackwell encyclopedia of social psychology* (pp. 505-509). Blackwell Publishers.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*(4), 349-354. https://doi.org/10.1037/h0047358

Cunningham, M. R., Steinberg, J., & Grev, R. (1980). Wanting to and having to help: Separate motivations for positive mood and guilt-induced helping. *Journal of Personality and Social Psychology, 38*(2), 181–192. https://doi.org/10.1037/0022-3514.38.2.181

Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior, 24*(6), 659–683. https://doi.org/10.1023/A:1005552203727

Darley, J. M., & Pittman, T. S. (2003). The psychology of compensatory and retributive justice. *Personality and Social Psychology Review, 7*(4), 324–336. https://doi.org/10.1207/S15327957PSPR0704_05

DePalma, M. T., Madey, S. F., Tillman, T. C., & Wheeler, J. (1999). Perceived patient responsibility and belief in a just world affect helping. *Basic and Applied Social Psychology, 21*(2), 131–137. https://doi.org/10.1207/15324839951036470

de Vignemont, F., & Singer, T. (2006). The empathetic brain: How, when, and why? *Trends in Cognitive Sciences, 10*(10), 435-441. https://doi.org/10.1016/j.tics.2006.08.008.

Dovidio, J. F. (1984). Helping behavior and altruism: An empirical and conceptual overview. In L. Berkowitz (Ed.), *Advances in experimental social psychology*: Vol. 17 (pp. 361–427). Academic Press.

Dovidio, J. F., Allen, J. L., & Schroeder, D. A. (1990). Specificity of empathy-induced helping: Evidence for altruistic motivation. *Journal of Personality and Social Psychology, 59*(2), 249–260. https://doi.org/10.1037/0022-3514.59.2.249

Dovidio, J. F., & Penner, L. A. (2003). Helping and altruism. In G. J. O. Fletcher, & M. S. Clark (Eds.), *Blackwell handbook of social psychology: Interpersonal processes* (pp. 162-195). Blackwell Publishers.

Dovidio, J. F., Piliavin, J. A., Gaertner, S. L., Schroeder, D. A., & Clark, R. D. III. (1991). The arousal: Cost-reward model and the process of intervention: A review of the evidence. In M. S. Clark (Ed.), *Prosocial behavior* (pp. 86–118). Sage Publications.

Duff, R. A. (1986). *Trials and punishments*. Cambridge University Press.

Dunning, D. (2011). Social motivation: Some introductory notes. In D. Dunning (Ed.), *Social motivation* (pp. 1-10). Psychology Press.

Eisenberg, N. (1991). Meta-analytic contributions to the literature on prosocial behavior. *Personality and Social Psychology Bulletin, 17*(3), 273–282. https://doi.org/10.1177/0146167291173007

Eisenberg, N. (1986). *Altruistic emotion, cognition, and behavior*. Erlbaum.

Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., & Reno, R. R. (1989). Relation of sympathy and personal distress to prosocial behavior: A multimethod study. *Journal of Personality and Social Psychology, 57*(1), 55–66. https://doi.org/10.1037/0022-3514.57.1.55

Eisenberg, N., & Miller, P. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin, 101*(1), 91–119.

Eisenberg, N., Miller, P. A., Schaller, M., Fabes, R. A., Fultz, J., Shell, R., & Shea, C. L. (1989). The role of sympathy and altruistic personality traits in helping: A reexamination. *Journal of Personality, 57*(1), 41–67. https://doi.org/10.1111/j.1467-6494.1989.tb00760.x

Emler, N., Renwick, S., & Malone, B. (1983). The relationship between moral reasoning and political orientation. *Journal of Personality and Social Psychology, 45*(5), 1073–1080. https://doi.org/10.1037/0022-3514.45.5.1073

Erkut, S., Jaquette, D. S., & Staub, E. (1981). Moral judgment-situation interaction as a basis for predicting prosocial behavior. *Journal of personality, 49*(1), 1-14. https://doi.org/10.1111/j.1467-6494.1981.tb00842.x

European Union. (2012). Charter of fundamental rights of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:12012P/TXT

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods, 41*(4), 1149-1160. https://doi.org/ 10.3758/BRM.41.4.1149

Finkel, N. J., Maloney, S. T., Valbuena, M. Z., & Groscup, J. (1996). Recidivism, proportionalism, and individualized punishment. *American Behavioral Scientist, 39*(4), 474-487. https://doi.org/10.1177/0002764296039004010

Finnis, J. (1972). The restoration of retribution. *Analysis, 32*(4), 131-135. https://doi.org/ 10.1093/analys/32.4.131

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review, 90*(4), 980–994. https://doi.org/10.1257/aer.90.4.980

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137-140. https://doi.org/10.1038/415137

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25*(2), 63–87. https://doi.org/10.1016/S1090-5138(04)00005-4

Finckenauer, J. O. (1988). Public support for the death penalty: Retribution as just deserts or retribution as revenge? *Justice Quarterly, 5*(1), 81-100. https://doi.org/10.1080/07418828800089621

Finnis, J. (1972). The restoration of retribution. *Analysis, 32*(4), 131–135. https://doi.org/10.1093/analys/32.4.131:31

Fishbach, A., & Ferguson, M. J. (2007). The goal construct in social psychology. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (pp. 490–515). The Guilford Press.

Foster, C. A., & Rusbult, C. E. (1999). Injustice and powerseeking. *Personality and Social Psychology Bulletin, 25*(7), 834–849. https://doi.org/10.1177/0146167299025007006

Frey, D. L., & Gaertner, S. L. (1986). Helping and the avoidance of inappropriate interracial behavior: A strategy that perpetuates a nonprejudiced self-image. *Journal of Personality and Social Psychology, 50*(6), 1083–1090. https://doi.org/10.1037/0022-3514.50.6.1083

Furnham, A. (2003). Belief in a just world: Research progress over the past decade. *Personality and Individual Differences, 34*(5), 795–817. https://doi.org/10.1016/S0191-8869(02)00072-7

Gaertner, S. L., & Dovidio, J. F. (1977). The subtlety of White racism, arousal, and helping behavior. *Journal of Personality and Social Psychology, 35*(10), 691–707. https://doi.org/10.1037/0022-3514.35.10.691

Gavrielides, T. (Ed.). (2016). *The psychology of restorative justice: Managing the power within.* Routledge.

Gentile, J. R. (2000). Learning, transfer of. In A. E. Kazdin (Ed.), *Encyclopedia of psychology*: Vol. 5. American Psychological Association.

Gerber, M. M., & Jackson, J. (2013). Retribution as revenge and retribution as just deserts. *Social Justice Research, 26*(1), 61-80. http://dx.doi.org/10.2139/ssrn.2136237

Gibbons, F. X. (1990). Self-attention and behavior: A review and theoretical update. In L. Berkowitz (Ed.), *Advances in experimental social psychology*: Vol. 23 (pp. 249-303). http://dx.doi.org/10.1016/S0065-2601(08)60321-4

Gintis H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology, 206*(2), 169-179. https://doi.org/10.1006/jtbi.2000.2111

Gollwitzer, P. M., & Brandstatter, V. (1996). Motivation. In A. S. N. Manstead & M. Hewstone (Eds.), *The blackwell encyclopedia of social psychology*. Blackwell Publishers.

Gouldner, A. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review, 25*, 161–178.

Greenwood, P. W. & Abrahamse, A. (1982). *Selective incapacitation*. RAND Corporation.

Gürerk, O., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science, 312,* 108-111. https://doi.org/10.1126/science.1123633

Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin, 32*(10), 1402–1413. https://doi.org/10.1177/0146167206291006

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Haist, M. (2009). Deterrence in a sea of "just deserts": Are utilitarian goals achievable in a world of "limiting retributivism"? *Journal of Criminal Law and Criminology, 99*(3), 789-822.

Hamilton, V. L., & Sanders, J. (1988). Punishment and the individual in the United States and Japan. *Law & Society Review, 22*(2), 301–328. https://doi.org/10.2307/3053438

Hart, H. L. A. (1968). *Freedom and responsibility: Essays in the philosophy of law*. Clarendon Press.

Hauser, M. D., Young, L., & Cushman, F. A. (2007). Reviving Rawls' linguistic analogy. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 2. The cognitive science of morality*. Oxford University Press.

Higgins, E. T., & Spiegel, S. (2004). Promotion and prevention strategies for self-regulation: A motivated cognition perspective. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 171–187). The Guilford Press.

Ho, R., ForsterLee, L., ForsterLee, R., & Crofts, N. (2002). Justice versus vengeance: Motives underlying punitive judgements. *Personality and Individual Differences, 33*(3), 365–378. https://doi.org/10.1016/S0191-8869(01)00161-1

Hoffman, M. L. (1981). Is altruism part of human nature? *Journal of Personality and Social Psychology, 40*(1), 121–137. https://doi.org/10.1037/0022-3514.40.1.121

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge University Press.

Hogarth, J. (1971). *Sentencing as a human process*. University of Toronto Press.

Hoover, C. W., Wood, E. E., & Knowles, E. S. (1983). Forms of social awareness and helping. *Journal of Experimental Social Psychology, 19*(6), 577–590. https://doi.org/10.1016/0022-1031(83)90017-3

Hudson, B. A. (2003). *Understanding justice: An introduction to ideas, perspectives and controversies in modern penal theory.* Open University Press.

Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology, 21*(3), 384–388. https://doi.org/10.1037/h0032317

Isen, A. M., & Simmonds, S. F. (1978). The effect of feeling good on a helping task that is incompatible with good mood. *Social Psychology, 41*(4), 346–349. https://doi.org/10.2307/3033588

Jones, R. A. (1970). Volunteering to help: The effects of choice, dependence and anticipated dependence. *Journal of Personality and Social Psychology, 14*(2), 121–129. https://doi.org/10.1037/h0028771

Kahn, A., & Tice, T. E. (1973). Returning a favor and retaliating harm: The effects of stated intentions and actual behavior. *Journal of Experimental Social Psychology, 9*(1), 43–56. https://doi.org/10.1016/0022-1031(73)90061-9

Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 31–53). Holt, Rinehart, & Winston.

Krebs, D. (1982). Psychological approaches to altruism: An evaluation. *Ethics, 92*(3), 447–458. https://doi.org/10.1086/292354

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* Appleton-Century-Crofts.

Leary, M. R. (2003). The self we know and the self we show: Self-esteem, self-presentation, and the maintenance of interpersonal relationships. In G. J. O. Fletcher, & M. S. Clark (Eds.), *Blackwell handbook of social psychology: Interpersonal processes* (pp. 457-477). Blackwell Publishers.

Leary, M. R., & Allen, A. B. (2010). Belonging motivation: Establishing, maintaining, and repairing relational value. In D. Dunning (Ed.), *Social motivation* (pp. 37-55). Psychology Press.

Leary, M. R., & Downs, D. L. (1995). Interpersonal functions of the self-esteem motive: The self-esteem system as a sociometer. In M. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 123–144). Plenum.

Lerner, M. J. (1980*). The belief in a just world: A fundamental delusion*. Plenum.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology, 75*(1), 5–18. https://doi.org/10.1037/0022-3514.75.1.5

Lewin, K. (1952). *Field theory in social science*. Tavistock.

Lobchuk, M. M., McClement, S. E., McPherson, C., & Cheang, M. (2008). Does blaming the patient with lung cancer affect the helping behavior of primary caregivers? *Oncology Nursing Forum, 35*(4), 681–689. https://doi.org/10.1188/08.ONF.681-689

Lönnqvist, J. E., Verkasalo, M., Wichardt, P. C., & Walkowitz, G. (2013). Personal values and prosocial behaviour in strategic interactions: Distinguishing value-expressive from value-ambivalent behaviours. *European Journal of Social Psychology, 43*(6), 554-569. https://doi.org/10.1002/ejsp.1976

Malle, B. F., & Dickert, S. (2007). Values. In R. F. Baumeister & K. D. Vohs (Eds.), *The encyclopedia of social psychology*. Sage.

Marci, C. D., & Orr, S. P. (2006). The effect of emotional Distance on psychophysiologic concordance and perceived empathy between patient and interviewer. *Applied Psychophysiology and Biofeedback, 31*(2), 115–128. https://doi.org/10.1007/s10484-006-9008-4

Marwell, G., & Schmitt, D. R. (1972). Cooperation in a three-person Prisoner's Dilemma. *Journal of Personality and Social Psychology, 21*(3), 376–383. https://doi.org/10.1037/h0032319

McCullough, M. E., & Tabak, B. A. (2010). Prosocial behavior. In R. F. Baumeister, & E. J. Finkel (Eds.), *Advanced social psychology: The state of the science* (pp. 263-302). Oxford University Press, Inc..

Miethe, T., & Lu, H. (2005). *Punishment: A comparative historical perspective*. Cambridge.

Midlarsky, E. (1991). Helping as coping. In M. S. Clark (Ed.), *Prosocial behavior* (pp. 238–264). Sage Publications, Inc.

Miller, D. T., & Vidmar, N. (1981). The social psychology of punishment reactions. In M. Lerner & S. Lerner (Eds.), *The justice motive in social behavior* (pp. 145–172). Academic Press.

Morris, H. (1981). Persons and punishment. In H. Gross & A. von Hirsch (Eds.), *Sentencing* (pp. 93-95). Oxford University Press.

Morris, H. (1994). A paternalistic theory of punishment. In A. Duff & D. Garland (Eds.), *A reader on punishment* (pp. 92-111). Oxford University Press.

Murphy, J. G. (1979). *Retribution, justice and therapy*. Springer.

Murray, H. A. (1938). *Explorations in personality*. Oxford University Press.

Nagin, D. S. (1998). Deterrence and incapacitation. In M. H. Tonry (Ed.), *The handbook of crime and punishment* (pp. 345–368). Oxford University Press.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*(4), 250–256. https://doi.org/10.1037/0022-3514.35.4.250

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*(5805), 1560–1563. https://doi.org/10.1126/science.1133755

Nowak, M., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature, 393*, 573–577. https://doi.org/10.1038/31225

Nowak, M., & Sigmund, K. (2005). The dynamics of indirect reciprocity. *Journal of Theoretical Biology, 194*(4), 561-574. https://doi.org/10.1006/jtbi.1998.0775.

Nozick, R. (1981). *Philosophical explanations*. Clarendon Press.

Oliner, S. P., & Oliner, P. M. (1988). *The altruistic personality: Rescuers of jews in Nazi Europe*. The Free Press.

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108*(1), 222-31. https://doi.org/10.1016/j.cognition.2007.12.003.

Organ, D. W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology, 48*(4), 775–802. https://doi.org/10.1111/j.1744-6570.1995.tb01781.x

Oswald, M. E., Hupfeld, J., Klug, S. C., & Gabriel, U. (2002). Lay-perspectives on criminal deviance, goals of punishment, and punitivity. *Social Justice Research, 15*(2), 85–98. https://doi.org/10.1023/A:1019928721720

Park, J. H., & Buunk, A. P. (2011). Interpersonal threats and automatic motives. In D. Dunning (Ed.), *Social motivation* (pp. 11-36). Psychology Press.

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology, 56*, 365-392. https://doi.org/10.1146/annurev.psych.56.091103.070141

Piazza, J., & Bering, J. M. (2008). Concerns about reputation via gossip promote generous allocations in an economic game. *Evolution and Human Behavior, 29*(3), 172–178. https://doi.org/10.1016/j.evolhumbehav.2007.12.002

Piliavin, J. A., Dovidio, J. F., Gaertner, S. L., & Clark, R. D., III. (1982). Responsive bystanders: The process of intervention. In V. J. Derlega & J. Grzelak (Eds.), *Cooperation and helping behaviour: Theories and research* (pp. 279–304). Academic Press.

Piliavin, J. A., & Piliavin, I. M. (1972). Effect of blood on reactions to a victim. *Journal of Personality and Social Psychology, 23*(3), 353–361. https://doi.org/10.1037/h0033166

Piliavin, I. M., Rodin, J., & Piliavin, J. A. (1969). Good samaritanism: An underground phenomenon. *Journal of Personality and Social Psychology, 13*(4), 289–299. https://doi.org/10.1037/h0028433

Pittman, T. S., & Zeigler, K. R. (2007). Basic human needs. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (pp. 473–489). The Guilford Press.

Prentice, D. (2000). Values. In A. E. Kazdin (Ed.), *Encyclopedia of psychology*: Vol. 8. American Psychological Association.

Price, M. E., Cosmides, L., & Tooby, J. (2002). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior, 23*(3), 203–231. https://doi.org/10.1016/S1090-5138(01)00093-9

Pruitt, D. G. (1968). Reciprocity and credit building in a laboratory dyad. *Journal of Personality and Social Psychology, 8*, 143-147. https://doi.org/10.1037/h0025323

Reber, A. S., & Reber, E. (Eds.). (2001). *The penguin dictionary of psychology*. Penguin Books.

Regan, D. T., Williams, M., & Sparling, S. (1972). Voluntary expiation of guilt: A field experiment. *Journal of Personality and Social Psychology, 24*(1), 42–45. https://doi.org/10.1037/h0033553

Reis, H. T., & Gruzen, J. (1976). On mediating equity, equality, and self-interest: The role of self-presentation in social exchange. *Journal of Experimental Social Psychology, 12*(5), 487–503. https://doi.org/10.1016/0022-1031(76)90079-2

Rapoport, A., & Chammah, A. M. (1965). Sex differences in factors contributing to the level of cooperation in the Prisoner's Dilemma game. *Journal of Personality and Social Psychology, 2*(6), 831–838. https://doi.org/10.1037/h0022678

Reykowski, J. (1982). Motivation of prosocial behavior. In V. J. Derlega & J. Grzelak (Eds.), *Cooperation and helping behavior: Theories and research* (pp. 352–375). New Academic Press.

Rind, B. & Strohmetz, D. (1999). Effect on restaurant tipping of a helpful message written on the back of customers' checks. *Journal of Applied Social Psychology, 29*(1), 139-144. https://doi.org/10.1111/j.1559-1816.1999.tb01378.x

Robinson, P. H. & Darley, J. M. (1997). The utility of desert. *Northwestern University Law Review, 91,* 453–499.

Sadurski, W. (1985). *Giving desert its due: Social justice and legal theory*. Reidel Publishing Company.

Salovey, P., Mayer, J. D., & Rosenhan, D. L. (1991). Mood and helping: Mood as a motivator of helping and helping as a regulator of mood. In M. S. Clark (Ed.), *Prosocial behavior* (pp. 215–237). Sage Publications, Inc.

Schmidt, P., & Witte, A. D. (1988). *Predicting recidivism using survival models*. Springer.

Schroeder, D. A., Penner, L. A., Dovidio, J. F., & Piliavin, J. A. (1995). *The psychology of helping and altruism: Problems and puzzles.* McGraw-Hill.

Schroeder, D. A., Dovidio, J. F., Sibicky, M. E., Matthews, L. L., & Allen, J. L. (1988). Empathic concern and helping behavior: Egoism or altruism? *Journal of Experimental Social Psychology, 24*(4), 333–353. https://doi.org/10.1016/0022-1031(88)90024-8

Schwartz, S. H. (1977). Normative influences on altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology*: Vol. 10 (pp. 221-279). Academic.

Schwartz, S. H. (1996). Values. In A. S. N. Manstead & M. Hewstone (Eds.), *The blackwell encyclopedia of social psychology*. Blackwell Publishers.

Schwartz, S. H., & Howard, J. A. (1982). Helping and cooperation: A self-based motivational model. In V. J. Derlega & J. Grzelak (Eds.), *Cooperation and helping behavior: Theories and research* (pp. 327–353). Academic Press.

Sedikides, C. (1993). Assessment, enhancement, and verification determinants of the self-evaluation process. *Journal of Personality and Social Psychology, 65*(2), 317–338. https://doi.org/10.1037/0022-3514.65.2.317

Seinen, I., & Schram, A. (2006). Social status and group norms: Indirect reciprocity in a repeated helping experiment. *European Economic Review, 50*, 581-602. https://doi.org/10.1016/j.euroecorev.2004.10.005.

Schaller, M., & Cialdini, R. B. (1988). The economics of empathic helping: Support for a mood management motive. *Journal of Experimental Social Psychology, 24*(2), 163–181. https://doi.org/10.1016/0022-1031(88)90019-4

Shaw, L. L., Batson, C. D., & Todd, R. M. (1994). Empathy avoidance: Forestalling feeling for another in order to escape the motivational consequences*. Journal of Personality and Social Psychology, 67*(5), 879–887. https://doi.org/10.1037/0022-3514.67.5.879

Shen, L. (2010). On a scale of state empathy during message processing. *Western Journal of Communication, 74*(5), 504-524. https://doi.org/10.1080/10570314.2010.512278

Sheposh, J. P., & Gallo, P. S. (1973). Asymmetry of payoff structure and cooperative behavior in the Prisoner's Dilemma game. *Journal of Conflict Resolution, 17*(2), 321–333. https://doi.org/10.1177/002200277301700208

Simpson, B., & Willer, R. (2015). Beyond altruism: Sociological foundations of cooperation and prosocial behavior. *Annual Review of Sociology, 41*, 43–63. https://doi.org/10.1146/annurev-soc-073014-112242

Skitka, L. J. (1999). Ideological and attributional boundaries on public compassion: Reactions to individuals and communities affected by a natural disaster. *Personality and Social Psychology Bulletin, 25*(7), 793–808. https://doi.org/10.1177/0146167299025007003

Skitka, L. J., & Tetlock, P. E. (1992). Allocating scarce resources: A contingency model of distributive justice. *Journal of Experimental Social Psychology, 28*(6), 491–522. https://doi.org/10.1016/0022-1031(92)90043-J

Skitka, L. J., & Tetlock, P. E. (1993). Providing public assistance: Cognitive and motivational processes underlying liberal and conservative policy preferences. *Journal of Personality and Social Psychology, 65*(6), 1205–1223. https://doi.org/10.1037/0022-3514.65.6.1205

Smith, K. D., Keating, J. P., & Stotland, E. (1989). Altruism reconsidered: The effect of denying feedback on a victim's status to empathic witnesses. *Journal of Personality and Social Psychology, 57*(4), 641–650. https://doi.org/10.1037/0022-3514.57.4.641

Sparks, P., & Durkin, K. (1987). Moral reasoning and political orientation: The context sensitivity of individual rights and democratic principles. *Journal of Personality and Social Psychology, 52*(5), 931–936. https://doi.org/10.1037/0022-3514.52.5.931

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193–210. https://doi.org/10.1037/0033-2909.103.2.193

Toi, M., & Batson, C. D. (1982). More evidence that empathy is a source of altruistic motivation. *Journal of Personality and Social Psychology, 43*(2), 281–292. https://doi.org/10.1037/0022-3514.43.2.281

Trivers, R. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology, 46*, 35–57. https://doi.org/10.1086/406755

United Nations. (2021). *The sustainable development goals report*. https://unstats.un.org/sdgs/report/2021/

Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Development, 81*(6), 1661-1669. https://doi.org/10.1111/j.1467-8624.2010.01500

VandenBos, G. R. (Ed.). (2015). *APA dictionary of psychology*. American Psychological Association.

Vidmar, N. (2001). Retribution and revenge. In J. Sanders & V. L. Hamilton (Eds.), *Handbook of justice research in law* (pp. 31–63). Kluwer Academic Publishers.

Vidmar, N., & Miller, D. (1980). Social psychological processes underlying attitudes toward legal punishment. *Law and Society Review, 14*, 565–602. https://doi.org/10.2307/3053193

Von Hirsch, A. (1976*). Doing justice: The choice of punishments: Report of the committee for the study of incarceration*. Hill and Wang.

Walster, E., Berscheid, E., & Walster, G. W. (1973). New directions in equity research. *Journal of Personality and Social Psychology, 25*(2), 151-156. https://doi.org/10.1037/h0033967

Warr, M., & Stafford, M. (1984). Public goals of punishment and support for the death penalty. *Journal of Research in Crime and Delinquency, 21*(2), 95-111. https://doi.org/10.1177/0022427884021002002

Weber, J. M., & Murnighan, J. K. (2008). Suckers or saviors? Consistent contributors in social dilemmas. *Journal of Personality and Social Psychology, 95*(6), 1340-1353. https://doi.org/10.1037/a0012454

Wedekind, C., & Braithwaite, V. (2002). The long-term effects of human generosity in indirect reciprocity. *Current Biology, 11*(12), 1012–15. https://doi.org/ 10.1016/s0960-9822(02)00890-4.

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in humans. *Science, 288*, 850-852. https://doi.org/10.1126/science.288.5467.850

Weiner, B. (1980). A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality and Social Psychology, 39*(2), 186–200. https://doi.org/10.1037/0022-3514.39.2.186

118

Weiner, B., Graham, S., & Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research, 10*(4), 431–452. https://doi.org/10.1007/BF02683293

Weiner, B., Perry, R. P., & Magnusson, J. (1988). An attributional analysis of reactions to stigmas. *Journal of Personality and Social Psychology, 55*(5), 738–748. https://doi.org/10.1037/0022-3514.55.5.738

Wenzel, M., & Thielmann, I. (2006). Why we punish in the name of justice: Just desert versus value restoration and the role of social identity. *Social Justice Research, 19*(4), 450–470. https://doi.org/10.1007/s11211-006-0028-2

Weyant, J. M. (1978). Effects of mood states, costs, and benefits on helping. *Journal of Personality and Social Psychology, 36*(10), 1169–1176. https://doi.org/10.1037/0022-3514.36.10.1169

Wilhelm, M. O., & Bekkers, R. (2010). Helping behavior, dispositional empathic concern, and the principle of care. *Social Psychology Quarterly, 73*(1), 11–32. https://doi.org/10.1177/0190272510361435

Williamson, G. M., & Clark, M. S. (1989). Providing help and desired relationship type as determinants of changes in moods and self-evaluations. *Journal of Personality and Social Psychology, 56*(5), 722–734. https://doi.org/10.1037/0022-3514.56.5.722

Williamson, G. M., Clark, M. S., Pegalis, L. J., & Behan, A. (1996). Affective consequences of refusing to help in communal and exchange relationships. *Personality and Social Psychology Bulletin, 22*(1), 34-47. https://doi.org/10.1177/0146167296221004

Yagi, Y., & Shimizu, T. (1996). Helping behavior following a failure experience. *Japanese Psychological Research, 38*(2), 53–65.

Yinon, Y., & Landau, M. O. (1987). On the reinforcing value of helping behavior in a positive mood. *Motivation and Emotion, 11*(1), 83–93. https://doi.org/10.1007/BF00992215

Zagefka, H., Noor, M., Brown, R., de Moura, G. R., & Hopthrow, T. (2011). Donating to disaster victims: Responses to natural and humanly caused events. *European Journal of Social Psychology, 41*(3), 353-363. https://doi.org/10.1002/ejsp.781

Appendices

Scenarios of the recordings: common part between conditions

ENTREVISTADOR: Na semana passada um acidente trágico atingiu a família Crespo, de Sintra. O Sr. e a Sra. Crespo não resistiram a uma colisão frontal numa estrada na freguesia do Cacém, no concelho de Sintra.

A família Crespo viveu em Sintra apenas por seis meses. Regressavam da sua residência anterior, no Cacém, onde tinham ido visitar uns amigos. O Sr. e a Sra. Crespo deixaram três filhos – Cátia, uma estudante finalista do ISCTE, em Lisboa; a Alice, de 11 anos; e o Marco, de 8 anos. O Tribunal de Família e Menores de Sintra deixou a Cátia a guarda temporária dos seus irmãos mais novos. Infelizmente, o Sr. e a Sra. Crespo não tinham seguro de vida, e os seus filhos ficaram com muito pouco dinheiro disponível. A Cátia procura desesperadamente manter a sua família unida, e terminar o seu curso. Ela espera obter a sua graduação no Verão de 2014, mas é confrontada com muitos problemas. Não tem dinheiro suficiente para comida e para a renda da sua casa. Necessita de uma pessoa que possa ficar com o seu irmão e irmã mais novos enquanto frequenta as aulas à noite, e precisa de dinheiro para transportes públicos para ir ao hipermercado e para a universidade, uma vez que não dispõe de automóvel. A Cátia está a tentar juntar dinheiro a partir de contribuições de particulares. Eu tive oportunidade da falar com a Cátia Crespo ontem. Ela descreve a sua situação nestas palavras.

CÁTIA: É um pesadelo. Eu ainda estou atordoada com isto. Eu sei que a vida tem de continuar. A coisa mais importante para mim neste momento é acabar o curso no tempo suposto. Eu tenho de estar em condições de arranjar um bom emprego, e sustentar o meu irmão mais novo e a minha irmã. A ajuda que temos tido até agora tem sido maravilhosa, mas ainda temos muito pela frente. E se não tivermos mais ajuda, vou ter mesmo que deixar a universidade e procurar um segundo trabalho, e isso assim vai ser pior, acho eu, porque toda a gente sabe que hoje em dia sem um curso superior não se consegue muito dinheiro. Se eu tiver de deixar a universidade, acho que vou ter de pôr as crianças numa instituição. E eu... eu não vou conseguir ter dinheiro suficiente para as sustentar...

In this moment participants listened to the manipulated content (see Appendix B).
Then followed:

ENTREVISTADOR: Estamos de facto perante uma situação complicada, que todos esperamos que seja rapidamente ultrapassada Cátia. Muito obrigado por a ter partilhado aqui connosco.

ENTREVISTADOR: Obterá a Cátia a ajuda que necessita? Difundimos esta história com a esperança de que alguns dos nossos ouvintes possam ajudar.

Se algum ouvinte desejar fazer uma contribuição de qualquer valor, por favor ligue para a [radio name censored effect] através do 214 144 140. O nosso programa chega ao fim, eu sou o Filipe Lamelas, e marcamos encontro para amanhã à mesma hora.

Appendix B

Scenarios of the recordings: manipulated content, varying between conditions

High harm and high history conditions:

ENTREVISTADOR: Não há dúvida que a Cátia se encontra numa situação muito delicada… nós temos conhecimento de que a Cátia sempre teve problemas com a sua equipa, no seu emprego, que certamente contribuíram para esta circunstância frágil/difícil que está a atravessar… quer falar-nos sobre isso?...

CÁTIA: Bem, sabe… eu sou uma pessoa um bocadinho vingativa.
(silêncio)
ENTREVISTADOR: (insiste) O que é que a Cátia quer dizer com isso?... Consegue dar-me um exemplo?...

CÁTIA: Sim, não há problema… o que aconteceu foi que, no meu local de trabalho, normalmente uma colega minha, que trabalha no gabinete ao lado do meu, costuma dar-me papelada para eu tratar, ou eu dou-lhe a ela, o que dá jeito, às vezes, quando temos muito que fazer.
Um dia essa minha colega pediu-me para preparar correio, para enviar a clientes da empresa – o que implica escrever os endereços nas cartas, criar etiquetas (para colar nos envelopes), imprimir tudo, e meter as cartas dentro dos envelopes – e eu passei uma tarde inteira a fazer isto.
Quando acabei levei-lhe os envelopes (para ela depois deixar no correio), e quando me viu, disse-me que se tinha distraído, e que acabou por não me avisar que o nosso chefe de serviço tinha mudado de ideias, e que queria que as cartas fossem enviadas por e-mail.
[In the Pretest and Study1 (hitting):] E quando ouvi isto dei-lhe um estalo na cara.
[In Study 2 (threatening):] E eu respondi-lhe que ela agora tinha de entregar as cartas mesmo assim, caso contrário eu contava ao nosso chefe de todas as vezes que ela saiu mais cedo do trabalho… e… que ela seria despedida depois disso.

(silêncio)
ENTREVISTADOR: E esta situação que descreveu é um exemplo de um problema habitual?

CÁTIA: Sim, é.

High harm and low history conditions:

ENTREVISTADOR: Não há dúvida que a Cátia se encontra numa situação muito delicada… nós temos conhecimento de que a Cátia passou recentemente por um episódio um pouco infeliz no seu trabalho, que certamente contribuiu para esta circunstância frágil/difícil que está a atravessar… quer falar-nos sobre isso?...

122

CÁTIA: Sim, não há problema… o que aconteceu foi que, no meu local de trabalho, normalmente uma colega minha, que trabalha no gabinete ao lado do meu, costuma dar-me papelada para eu tratar, ou eu dou-lhe a ela, o que dá jeito, às vezes, quando temos muito que fazer.

Um dia essa minha colega pediu-me para preparar correio, para enviar a clientes da empresa – o que implica escrever os endereços nas cartas, criar etiquetas (para colar nos envelopes), imprimir tudo, e meter as cartas dentro dos envelopes – e eu passei uma tarde inteira a fazer isto.

Quando acabei levei-lhe os envelopes (para ela depois deixar no correio), e quando me viu, disse-me que se tinha distraído, e que acabou por não me avisar que o nosso chefe de serviço tinha mudado de ideias, e que queria que as cartas fossem enviadas por e-mail.

[In the Pretest and Study1 (hitting):] E quando ouvi isto dei-lhe um estalo na cara.

[In Study 2 (threatening):] E eu respondi-lhe que ela agora tinha de entregar as cartas mesmo assim, caso contrário eu contava ao nosso chefe de todas as vezes que ela saiu mais cedo do trabalho… e… que ela seria despedida depois disso.

Low harm and high history conditions:

ENTREVISTADOR: Não há dúvida que a Cátia se encontra numa situação muito delicada… nós temos conhecimento de que a Cátia sempre teve problemas com a sua equipa, no seu emprego, que certamente contribuíram para esta circunstância frágil/difícil que está a atravessar… quer falar-nos sobre isso?...

CÁTIA: Bem, sabe… eu sou uma pessoa um bocadinho vingativa.
(silêncio)
ENTREVISTADOR: (insiste) O que é que a Cátia quer dizer com isso?... Consegue dar-me um exemplo?...

CÁTIA: Sim, não há problema… o que aconteceu foi que, no meu local de trabalho, eu normalmente ajudo os meus colegas a tratar de papelada, e eles ajudam-me a mim, quando preciso.

Um dia uma colega minha, que trabalha no gabinete ao lado do meu, pediu-me para a ajudar nas suas tarefas para aquele dia, porque achava que não teria tempo para fazer tudo. Pediu-me para preparar correio para ser enviado a clientes da empresa – o que implica escrever os endereços, adaptar as cartas para cada cliente, criar etiquetas (para colar nos envelopes), imprimir tudo, e meter as cartas dentro dos envelopes – e passei uma manhã inteira a fazer isto.

Quando acabei levei-lhe os envelopes, e quando me viu, disse-me que se tinha distraído, e que acabou por não me avisar que o nosso chefe de serviço tinha mudado de ideias, e que queria que as cartas fossem enviadas por e-mail – mas que gostaria de usar os ficheiros do Word, das cartas, porque isso lhe pouparia muito tempo.

E eu recusei-me a dar-lhe os ficheiros.

123

(silêncio)

ENTREVISTADOR: E esta situação que descreveu é um exemplo de um problema habitual?

CÁTIA: Sim, é.

Low harm and low history conditions:

ENTREVISTADOR: Não há dúvida que a Cátia se encontra numa situação muito delicada… nós temos conhecimento de que a Cátia passou recentemente por um episódio um pouco infeliz no seu trabalho, que certamente contribuiu para esta circunstância frágil/difícil que está a atravessar… quer falar-nos sobre isso?...

CÁTIA: Sim, não há problema… o que aconteceu foi que, no meu local de trabalho, eu normalmente ajudo os meus colegas a tratar de papelada, e eles ajudam-me a mim, quando preciso.
Um dia uma colega minha, que trabalha no gabinete ao lado do meu, pediu-me para a ajudar nas suas tarefas para aquele dia, porque achava que não teria tempo para fazer tudo. Pediu-me para preparar correio para ser enviado a clientes da empresa – o que implica escrever os endereços, adaptar as cartas para cada cliente, criar etiquetas (para colar nos envelopes), imprimir tudo, e meter as cartas dentro dos envelopes – e passei uma manhã inteira a fazer isto.
Quando acabei levei-lhe os envelopes, e quando me viu, disse-me que se tinha distraído, e que acabou por não me avisar que o nosso chefe de serviço tinha mudado de ideias, e que queria que as cartas fossem enviadas por e-mail – mas que gostaria de usar os ficheiros do Word, das cartas, porque isso lhe pouparia muito tempo.
E eu recusei-me a dar-lhe os ficheiros.
Depois a minha colega acabou por ser repreendida, por não ter enviado todos os e-mails a tempo.