Intentions to Steal and the Commitment Problem. The Role of Moral Emotions and Self-Serving Justifications

Ann De Buck[1,2,3] (iD) and Lieven J. R. Pauwels[1,2,3]

Abstract

This study focuses on determinants underlying young persons" self-reported intentions to steal a small amount of money. From an evolutionary standpoint, theft is a frequency-dependent strategy that may have been favored because it gave individuals a reproductively-relevant advantage in the competition for scarce resources. Although human groups do not tolerate the act of stealing, theft is still very common. Our study is rooted in Robert Frank's theory of the moral commitment problem. Moral emotions such as anticipated guilt are devices designed by evolutionary forces to motivate cooperative behavior in situations entailing a commitment problem. However, the anticipation of guilt feelings can be circumvented by self-serving justifications, therefore increasing the likelihood to steal. A large region-wide sample of adolescents (N = 3694) is used to analyze whether anticipated moral guilt and self-serving justifications mediate the effects of empathy, fear sensitivity, and perceived peer disapproval in their relationship to intentions to steal. Several propositions are tested in a latent variable model within the framework of SEM. Visual scenarios depicting an opportunity to take a small amount of money from a stranger are used to elicit participants" self-reported intentions to steal. Results suggest that empathic concern and empathic perspective-taking, perceived peer disapproval, and fear sensitivity affect the likelihood of theft by influencing anticipated guilt and self-serving justifications that, in turn, respectively reduce and promote the likelihood of theft.

Introduction

Humans are an interdependent social species (Tomasello, 2016). Given that our ancestors faced many life-threatening challenges (e.g., food scarcity, and the threat of predators), individuals who stayed nearby and cooperated gained significant survival and reproductive advantages over solitary individuals (Campbell, 1982; Richerson & Boyd, 1998; Turchin, 2015). However, in social systems, individuals inevitably encounter opportunities to exploit others and reap the benefits of cooperation without paying their share of the costs (Krebs, 2011). Living in close cooperative groups is costly because cooperation is vulnerable to cheating, it makes group members, among other things, susceptible to theft (Neuberg & DeScioli, 2016). Theft is universal in human cultures and is also observed among other species, strongly suggesting a biological and

evolutionary origin (Kanazawa, 2008) A cost-inflicting strategy such as theft may have been favored by natural selection because it gave individuals a reproductively-relevant advantage in competitions for scarce resources (Duntley & Shackelford, 2008). Opportunities to cheat (to steal) are ubiquitous in social groups, however, cooperative human groups do not tolerate the act of stealing and they punish individuals who take

[1] Faculty of Law and Criminology, Ghent University, Ghent, Belgium
[2] Department of Criminology, Criminal Law and Social Law, Ghent, Belgium
[3] Institute for International Research on Criminal Policy (IRCP), Ghent, Belgium

Corresponding Author:
Ann De Buck, Faculty of Law and Criminology, Ghent University, Universiteitstraat 4, 9000 Ghent, Belgium.
Email: ann.debuck@ugent.be

personal property that belongs to someone else (Boehm, 2012). Notwithstanding, the lifetime prevalence of stealing appears fairly high. For example, the National Association of Shoplifting and Theft Addiction Specialists estimates that 10% of the US's population are shoplifters.[1]

The commitment problem captures a fundamental problem of human sociality, that is, a problem of confluences and conflicts of interest. Confluences of interest because people help each other to reap more optimal outcomes than they could by living a solitary life, and conflicts of interest, because it is in each individual's interest to induce others to give more and take less than they do (Alexander, 1987; Dawes, 1980; Hardin, 1968; Krebs, 2008; Rawls, 1999). In game theory, the commitment problem arises if one would be better off resisting the temptation to cheat, thereby, fostering the benefits of pro-social behavior, instead of opportunistically pursuing one's self-interest without concern for others (Frank, 1988, 2001). How can people facing commitment problems be motivated to reap the benefits of delayed rewards? Frank (1988) proposed moral emotions as commitment devices. His proposition is a theory of the function of moral emotions, such as the anticipation of guilt, that motivates to refrain from selfishly pursuing immediate rewards, therefore inhibiting antisocial behavior. Many empirical results support this notion (e.g., Olthof, 2012; Pauwels & Svensson, 2015; Rebellon et al., 2016). However, guilt feelings can be circumvented or reduced by self-serving justifications that result from a biased view regarding the propriety of one's behavior (Shalvi et al., 2015; Sykes & Matza, 1957). Self-serving justifications are thought of as cognitive mechanisms that operate by restructuring one's wrongdoings and consequences, and reducing the guilt that inhibits antisocial behavior, thereby facilitating such behavior (Bandura, 1991). Empirical results support this notion (e.g., Agnew, 1994; Copes, 2003; Ribeaud & Eisner, 2010; Siebert & Stewart, 2019; Walters et al., 2021; also see a meta-analytic review (Walters, 2020)). Anticipated guilt and self-serving justifications can be thought of as proximate mechanisms underlying antisocial behavior such as theft.

In this study, we seek to contribute to knowledge about facets of morality as problem-solving devices by investigating individual variations in antecedents of intentions to steal. We propose an integrated conceptual framework in which anticipated guilt and self-serving justifications, which are operative in the individual minds *at the time* of decision-making, are proximate factors that respectively inhibit and promote the likelihood of theft. Our model also incorporates individual characteristics and perceptions as more distal factors underlying the proximate factors. Specifically, we hypothesize that individual differences in empathic concern and perspective-taking, perceived peer disapproval, and fear sensitivity relate to how individuals perceive themselves, others (peers), and the environment. We propose that these individual differences are associated with anticipated guilt and self-serving justifications that, in turn, have respectively inhibitory and promoting influences on making some individuals more likely to steal. Two visual scenarios depicting an opportunity to steal a small amount of money from a stranger are used to elicit research participants" self-reported behavioral tendencies and responses to the explanatory variables.

The Commitment Problem: The Roles of Anticipated Guilt and Self-Serving Justifications

The Commitment Problem

Humans live in uniquely cooperative societies. People cooperate with unrelated individuals in every human society, compared to other primate species whose cooperation is limited to relatives and small groups of reciprocating partners (Boyd, 2006). In addition, human social life is regulated by moral systems which is essentially a society with agreements about what is permitted and what is not, agreements that are enforced by third parties (although not always perfectly) (Alexander, 1985). Solitary organisms do not need cooperation. The way we understand cooperation here is how Tomasello and Vaish (2013) have put it: "…*What cooperation requires is that individuals either suppress their self-interest or that individuals equate it with the interests of others*" (p. 231). It is difficult to exaggerate the significance of conflicts of interest (Alexander, 1979, 1987). Human society is filled with individuals who are following what they perceive to be their interests. That is, according to Alexander (1987), the most general principle of human behavior (p. 34). From an evolutionary biological viewpoint, conflicts of interest arise out of the fact that humans are genetically unique (Alexander, 1987). The notion that humans have an individualized set of genes is not trivial to understand that humans have evolved to have *individually separate interests* (Alexander, 1979, p. 17), and to strive, in relation to one another, for the same things which not everyone can possess or possess equally because resources are finite (Alexander, 1987).

Human sociality comes with a fundamental problem. Cooperating with others had benefits in contexts in which working together paid off and was more beneficial than one alone could achieve, such as cooperative hunting, sharing food, helping in times of danger (e.g., defense against predation or attacks from rival groups, sharing knowledge, building shelter (Trivers, 2002)). However, conflicts over scarce resources (e.g., food, mates, territory…) between individuals would have posed recurrent problems over human evolutionary history (Buss & Shackelford, 1997). There would have been significant selection pressures in favor of an adaptive behavioral repertoire favoring cooperation with others when it paid to do so (in terms of contribution to an individual's reproductive success), and in favor of behaviors for out-competing others in the context of recurrent conflicts between individuals (Duntley & Shackelford, 2008). Behavior that societies define as selfish[1], deviant, or criminal (e.g., expropriation of goods or services) provides expressions of a behavioral repertoire that may have been favored over evolutionary time when they gave individuals a reproductive advantage in social and ecological environments in which they developed (Cohen & Machalek, 1988; but also

see, Durrant & Ward, 2015; Raine, 1997; Walsh & Jorgensen, 2018).

In formal game theory, cooperation yields the best outcome for all if all work in unison, but each individual faces the temptation to step back and let others pay the cost of cooperation while enjoying the benefits that cooperation brings about. That is the original definition of a social dilemma (Dawes, 1980). The same idea underlies the commitment model of the economist Frank (1988) who refers to commitment problems as an important class of recurring problems that people confront in their daily social interactions. Opportunities to cheat (to lie or to steal) are ubiquitous in every culture (Boehm, 2012), hence humans'' adaptive ability to detect cheaters (Buss, 2019; Cosmides & Tooby, 1992, 2008). Where such opportunities exist and detection of cheating is impossible or small, there are always opportunistic persons who are readily bound to exploit them. Having the means to make binding commitments not to cheat (to lie or to steal) would benefit everyone (Frank, 1988). Given the ubiquity of commitment problems constraining people's ability to reap the benefits of cooperation, the question arises of how they are dealt with. Frank (1988) proposes moral emotions, like guilt, as devices by which cooperation may be induced in people facing commitment problems. Guilt is thought to motivate people to control their impulses to pursue immediate rewards, and therefore to inhibit uncooperative behavior. However, people vary in their tendency to feel guilt (Tangney, 1990), and guilt feelings can be reduced or circumvented by self-serving justifications (Sykes & Matza, 1957; Walters, 2020).

The Role of Anticipated Guilt

Guilt is understood as an interpersonal phenomenon that happens in social interactions, as much as it is an intra-psychic emotion (Baumeister et al., 1994). Attributional theories, on the one hand, focus on the intrapersonal nature of guilt as a subjective painful emotion of feeling guilty for having broken the rules of society (Katchadourian, 2010). During the socialization by parents and extra-parental conspecifics, people internalize the rules of conduct that translate to the values and rules of their social groups (Gintis, 2003). Individuals connect emotionally with those rules so that they feel good about following them and uneasy about breaking them (Boehm, 2012). Behavior that is not in line with internalized standards may result in self-evaluation and self-censure, such as feelings of guilt (Tangney & Dearing, 2003), and push in the direction of social conformity (Tangney & Stuewig, 2004). Adaptationist perspectives, on the other hand, focus on the interpersonal nature of guilt as an evolved information-processing adaptation that helps to navigate opportunities and threats in the social world related to the social valuation of interaction partners (Cosmides & Tooby, 2000). Guilt is hypothesized to be activated when an individual perceives she has placed too low a weight on another person's welfare (Tooby & Cosmides, 2008), therefore, guilt motivates to benefit the victim and to repair a valuable relationship (Sznycer, 2019).

Robert Frank's commitment model (1988) emphasizes the role of guilt as a proximate psychological reward mechanism that governs behavior. Frank argues that guilt is a problem-solving device because feeling guilty or the anticipation of guilt often predisposes people to behave in ways that may seem contrary to narrow psychological self-interest.[2] The emotion of guilt can act as a constraint to a self-serving course of action. Guilt can cause a person who feels bad (guilty) when he cheats, to behave honestly even when she knows she could get away with cheating. A person capable of strong guilt feelings will not cheat even when it is in her material interests to do so. She simply doesn't want to cheat because of her aversion to the painful feelings of guilt. That person's painful feelings commit her to behave honestly in the face of a golden opportunity to cheat. Because a commitment is a promise, those who break a promise come across as untrustworthy, and those who keep their promise come across as trustworthy (Krebs, 2011). Emotional predispositions in addition to being known as a person who experiences guilt enable one to make commitments that would otherwise not be credible in ventures that require trust. The proposed function of guilt is to motivate people to behave in ways that are contrary to narrow interests in the short term, thus promoting prosocial behavior, and being committed to forgoing immediate selfishness to reap the long-term reward of cooperation, can have a reputational advantage in the long run. Reputation is important in the life of humans (and in the lives of other social organisms) because it is thought to be, at least partly, responsible for human cooperation among nonkin (Alexander, 1987; Barclay, 2016). The fact that guilty feelings are expected after a perceived wrong deed makes the anticipation of feeling guilty a threat of punishment. Anticipated guilt can act as a strong disincentive to give in to the temptation to steal (Baumeister et al., 2007).

The Role of Self-Serving Justifications

Well-internalized moral values and rules do not make people socially perfect, but rather ambivalent conformists (Boehm, 2012). Alexander (1987) defined the evolutionary conscience as a means to being a moral person but at the same time as a means to profit from a tempting but socially disapproved behavior and to strategize about it (Boehm, 2012). The tendency to have biased views regarding the propriety of one's conduct was termed *self-deception* by Adam Smith in his *Theory of Moral Sentiments* (Smith, 1817(1759)). Self-deception makes it hard to see one's conduct from the outlook of an impartial spectator (p. 81). Self-deception is also the concept used by Trivers (2002), who argued that self-deceit could easily be induced by our social nature because of its adaptive function: that is, self-deception was selected to deceive ourselves in the service of improved deception of others. The cost, on the other hand, is an active misrepresentation of reality to oneself. Trivers advances multiple sources of self-deception such as the construction of biased social theory and fictitious narratives of intentions. In Trivers'' view, social theory consists of a complex array of facts organized to be biased in favor of the speaker. Fictitious narratives of intentions are thoughts, and verbalizations that may act

as rationalizations for what we are doing, immediately available verbally should we be challenged by others who hide true intentions. We may be aware of the fact that stealing is not OK but at the same time deny it (Trivers, 2002). This idea is similar to the concept of *moralistic modules* that cause people to condemn conduct X while other modules cause doing X, leading to hypocrisy (Kurzban, 2011). Self-serving justifications are thought to circumvent anticipated guilt before transgressing and thus, people can act unconstrained by the self-regulatory effect of guilt (Sykes & Matza, 1957). For instance, anticipated guilt can be reduced if people believe that a victim is responsible for their wrongdoings or the consequences (displacement of responsibility; Bandura, 1991). Empirical evidence suggests that self-serving justifications increase the likelihood of antisocial behavior; also, evidence points to an inverse relationship with anticipated guilt (Ring & Kavussanu, 2018; Stanger et al., 2013). In our study, we treat anticipated guilt and self-serving justifications as two related proximate variables underlying intentions to steal. Hence, we hypothesize that:

H1: Anticipated guilt decreases the likelihood to steal, whereas self-serving justifications increase the likelihood to steal.

The Roles of Empathy, Perceived Peer Disapproval, and Fear Sensitivity

So far, we focused on the roles of anticipated guilt and self-serving justification as proximate psychological factors accounting for individual differences in intentions to steal. However, in any given situation people bring their unique set of personality characteristics, preferences, previous experiences, current goals, and everything they care about, which shape their perceptions of the situation. These factors include, among other things, personal moral standards, empathic tendencies, and fear sensitivity that influence their appraisal of features of the environment, such as opportunities, that are significant for them (Moors et al., 2013). These individual differences in shaping perception of and the degree to which one is tempted by opportunities may explain variation in responses to the same objective situation (Walsh, 2010). Specifically, in the following paragraph, we pay attention to antecedent traits, i.e., empathy and fear sensitivity, and perceptions of peer disapproval.

The Roles of Empathic Concern and Empathic Perspective-Taking

Adam Smith (1817[1759]), and later Darwin (1981[1871]) viewed dispositions to compassion and sympathy as basic social sentiments of human nature. Although Smith did not use the word empathy, he put forth the thesis that compassion is the sentiment that allows us to feel what others feel, that we bring to ourselves by imaginatively putting ourselves in their position. The terms sympathy (or compassion) and empathy are often used interchangeably. Both are closely linked and yet they are different. In Eisenberg's view (2000), pure empathy is not other-oriented but an affective response that is similar to what another

person is feeling. With further cognitive processing to differentiate between one's own and other's internal states, the affective response usually turns into sympathy or personal distress, or both. Sympathy consists of feelings of sorrow or concern for the other, which is not the same as feeling as the other person feels. Sympathy stems from the comprehension of the other's emotional condition (Eisenberg, 2000). Batson (2011) argued that sympathy or empathic concern is associated with other-oriented motivation fostering altruism. According to Bloom (2013), empathy (feeling what another person feels) motivates compassion, that is, feeling and acting kindly toward another without necessarily feeling how the other person feels. Caring and compassionate feelings toward another person are what contemporary psychologists call the affective or emotional component of empathy, whereas the ability to understand another person's perspective is the cognitive component of empathy (Maibom, 2014). Haidt (2003) links empathic concern to moral emotions. He defines moral emotions as those emotions which focus on others, that are elicited by stimuli that do not directly touch the self (like sympathy most strongly felt for one's kin), and that result in an outcome that benefits others or the social order. Hoffman (2000) posits a developmental model in which empathic feelings for someone else's distress, combined with the awareness of being personally responsible for his/her distress, results in feelings of guilt that in turn motivate prosocial behavior and inhibits antisocial behavior. Martinez et al. (2014) proposed and empirically tested an empathic pathway, termed the *empathic triad*, in which empathic perspective-taking serves as an antecedent cognitive process to feelings of empathic concern. If a person can cognitively connect to and imagine another person's mental state via empathic perspective-taking, one can become emotionally connected via the experience of empathic concern. Empathic concern, in turn, feeds into other emotional other-regarding capacities that similarly implicate concern for others" welfare, namely the tendency to feel guilt should one perceive to be the source of harm to another (Hoffman, 2000). Activation of felt concern for the feelings of another person may give rise to anticipated feelings of guilt if one faces the option to harm another person. The proposition is that this empathic triad (perspective-taking—empathic concern—guilt-proneness) would diminish transgressions. Overall, Martinez and colleagues (2014) found support for the notion that perspective-taking propensities negatively predict criminal behavior (diverse measures were used) via a psychologically proximal link with empathic concern and guilt-proneness. Increased perspective-taking and having sympathy (empathic concern) for others may activate the restraining influence of anticipated guilt over failures to cooperate (Leith & Baumeister, 1998). In addition, if people are more likely to blame the victim, thus resorting to self-serving justifications, then it is likely that they are less concerned about the welfare of the other person (Bandura, 1991, 2017; Detert et al., 2008). Hence, we would expect that:

H2: Empathic concern and empathic perspective-taking are positively related to anticipated guilt, and negatively related to self-serving justifications.

The Role of Perceived Peer Disapproval

The second factor to consider is the role of perceived peer disapproval. In addition to sympathy, both Adam Smith (1817[1759]), and Darwin (1981[1871]) considered a desire to avoid other people's disapproval as a social sentiment. Smith saw the principle of approval and disapproval as an instance of a human natural desire for sympathy. Having best friends is one of the most valuable commodities social creatures like humans can have. Best friends are likely to help in times of need, making best friendship a scarce resource (Tooby & Cosmides, 1996). Selection could have led to the evolution of cognitive mechanisms that motivate individuals to form and maintain such beneficial social relationships (Lewis et al., 2015). A strong desire to belong, and to be valued by others, is one of the most universal and powerful human drives (Baumeister, 2011; Baumeister & Leary, 1995). Consequently, caring about how friends value us, is part of our psychology (Raihani, 2021). The way our friends value us depends on how they think of us. How we behave, influences their judgments of us, suggesting that the psychology that motivates our speech and behavior should be designed to generate as positive a view as possible (Kurzban, 2011). Perceptions of approval/disapproval are particularly important during adolescence and young adulthood. Adolescence is a period of heightened susceptibility/sensitivity to environmental cues (positive and negative), especially in the presence of peers. Relative to children and adults, adolescents show an increased responsivity to reward cues, such as money, status, and peer acceptance (Sapolsky, 2017; Somerville, 2013). Advances in developmental neuroscience suggest that this sensitivity is related to changes in the dopamine-rich regions of the brain (Casey, 2015; Steinberg, 2008). Thus, what we do, influences how best friends judge us as valued persons. Social disapproval can coerce people into forgoing the temptation of short-term benefits rather than taking what belongs to someone else because non-cooperators may be judged as unreliable and undesirable as a friend (Raihani, 2021). If a person thinks that their best friend would disapprove of stealing, and value him/her less as a consequence, we would expect that person to refrain from stealing and be more likely to anticipate guilt and less inclined to resort to self-serving justifications.

H3: Perceived peer disapproval is positively related to anticipated guilt and negatively related to self-serving justifications.

The Role of Fear Sensitivity

The last factor to consider is the role of fear sensitivity. Fear is a basic emotion that has evolved because it served survival and reproductive functions (Panksepp & Watt, 2011). Fear is a universal and strong core emotion with distinctive physiological manifestations (Ekman, 1999). Boehm (2012) argues that our moral origins are rooted in "old ancestral" fear-based mechanisms (e.g., fear of punishment) that have been supplemented by newly evolved traits such as sophisticated perspective-taking, internalization of rules, a sense of shame, and a moralistic

judgment of oneself (also see Walsh, 2018). Fear-proneness is found to be implicated directly in children's moral conduct and appears to be a predictor of guilt, an emotion that develops later (Baker et al., 2012). Research indicates that fear-proneness is inversely associated with children's tendency to violate rules and that this association appears to be mediated by guilt (Kochanska et al., 2002). Anticipating emitting behavior that has been punished in the past and therefore could be punished, can evoke fear (Krebs, 2011). Fear of punishment is found to be inversely related to self-serving justifications (Čermák & Blatný, 1995). Hence, we expect that:

H4: Fear sensitivity is positively related to anticipated guilt and negatively related to self-serving justifications.

The Present Study

The present study focuses on the intermediary role of anticipated guilt and self-serving justifications, as proximate psychological factors accounting for individual variability in intentions to steal. Further, we pay attention to antecedent traits (empathy, fear sensitivity) and perceived disapproval by peers as more distal factors. We hypothesize that anticipated guilt and self-serving justifications are immediate predictors with respectively, negative and positive effects on intentions to steal (direct effects). We further propose that participants high in empathic traits (concern and perspective taking), high in perceptions of peer disapproval, and fear sensitivity experience higher levels of anticipated guilt and lower levels of self-serving justifications in a way that mediates the likelihood of theft (indirect effects). Figure 1 represents the conceptual model which we confront with our survey data. We test our hypotheses using structural equation models (SEMs).

Methods

Participants and Procedure

Data were gathered, through an online survey, of adolescents and young adults in a region-wide sample in the Dutch-speaking part of Belgium in 2019. An invitation to participate in the online survey was sent by email to all three hundred Flemish city councils and their youth associations and school boards asking for their cooperation. The survey was advertised as a broad study into choices young people make in social and moral situations. Participants could participate in a raffle, via a separate link. A convenience sample of 3694 participants, aged between 12 and 25 years old, completely finished the survey (1188 males (32.2%) and 2506 females (67.8%) (age_males : $M = 17.06$, $SD = 3.09$; age_females: $M = 17.65$, $SD = 3.18$)). To get an insight into the effects of potential biases due to the oversampling of females, we conducted separate analyses for males and females. The subgroup analyses, based on biological sex, showed the same pattern of results which suggests greater confidence that the oversampling of females does not bias the overall results.
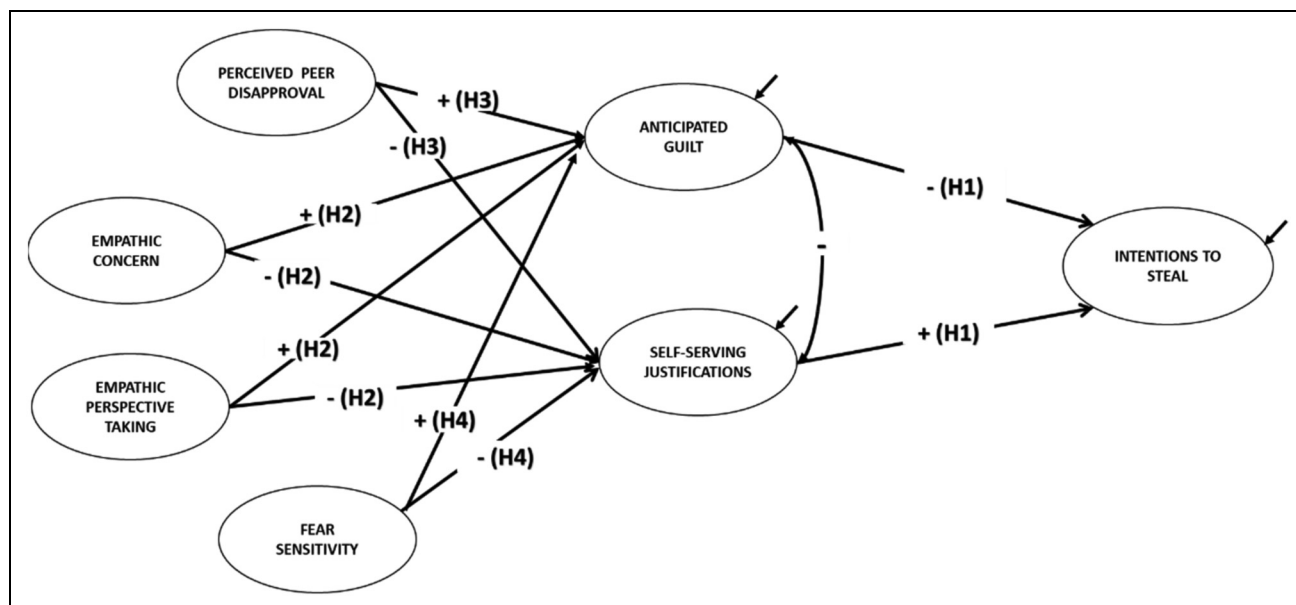
Figure 1. The conceptual model predicting intentions to steal, anticipated guilt, and self-serving justifications.

Visual Scenarios

A visual scenario design comprising two different scenarios was developed to test our conceptual model. Two hypothetical scenarios were filmed from the viewpoint of the perceiver.[3] This way we were able to put participants *in the shoes* of a would-be offender. The visual scenarios displayed a potentially tempting situation to steal a small amount of money (€50)[4]. One scenario took place at an ATM, another in a coffeehouse. Each participant viewed two scenarios and was asked to imagine him/herself in the situation and to answer several questions about it. The scenario in a coffee shop is displayed as follows (for a description of the ATM scenario, see the online supplemental appendices):

> The main character is sitting in a coffee shop somewhere in town, drinking coffee and reading a magazine. A woman passing by stops at the table and says the following to the main character: *"Here, this was lying on the ground"* and then she puts a €50 note on the table in front of the main character. She then walks out of the coffee shop. The main character takes the €50 note and puts it in his jacket. One moment later, a young man enters the coffee shop. He appears to be searching for something, walks up to the table where the main character is sitting, and asks: *"Sorry, I just sat here and I lost a 50 euro note. Did you find it?"*

Measurement

Each scenario was followed by several indicators measuring the likelihood of theft (dependent variable), anticipated guilt, fear sensitivity, and perceived peer disapproval. We computed a full latent variables model by aggregating participants" responses to the two scenarios for each theoretical concept. Aggregated responses to the scenarios reduce error variance

and ensure more reliable and valid measures (see also van Gelder & de Vries, 2014). Scale items, factor loadings, and reliability analyses of each construct can be found in the supplemental appendices.

Dependent variable

Intentions to steal is the dependent variable. After watching each scenario, participants were asked to give their estimate of the likelihood that they would keep the money on a seven-point Likert scale: *"How likely is it that you would keep the €50?"* (very unlikely—very likely). In addition, a second indicator probed participants" estimate of the likelihood of stealing €200 instead of €50. The latent outcome variable was a measure consisting of four estimate indicators (two per scenario), measuring participants" uncooperative behavioral tendencies, with a good alpha reliability of.83. 324 participants (9,2%) consistently reported they considered it likely to very likely they would keep the money. 1230 participants (35,5%) consistently reported that it would be very unlikely that they would keep the money.

Intermediairy Variables

Anticipated guilt is captured using the *State Shame and Guilt Scale* (SSGS; Marschall et al., 1994). The scale is composed of brief phenomenological descriptions of guilt (five items) experiences. For present purposes, we adapted the descriptions of SSGS into a measure of *anticipated* guilt, expressed *at the moment* of decision-making. After watching each video, participants are asked to imagine *they had decided to keep the money* (€50), *irrespective of their answers to the previous questions inquiring about their intentions to steal*, and how they would *feel*. Example items are "*I would feel remorse, regret; I*

would feel tension about something I have done; I would not stop thinking about something bad I have done" (rated on a five-point Likert scale: 1 = "fully agree" to 5 = "fully disagree"). Responses were reverse coded, and high scores on the indicators indicate high levels of anticipated guilt. A latent variable was constructed with ten indicators (five per scenario) with a good alpha reliability of.93.

Self-serving justifications are measured using *two conduct-specific indicators*. Two questions were asked *"Stealing a small amount of money is OK when you consider that there are others who steal a lot of money"*, and *"If people are careless about where they leave their things, it is their fault they get stolen"*. Both indicators are measured using a seven-point Likert scale ranging from 1 (fully disagree) to 7 (fully agree). High scores on the indicators imply high levels of self-serving justifications with a moderate alpha reliability of.54.

Independent Variables

Empathic concern and empathic perspective-taking are independent variables that are assessed using the Dutch version of the Interpersonal Reactivity Index[5] (IRI; Davis, 1983, 1994). The IRI is a self-report multidimensional individual difference measure of empathic dispositions. In this study, we used two subscales: (1) the *Perspective Taking* scale (EP) contains seven items that assess the tendency to spontaneously adopt the psychological point of view of others and to see things from their point of view (e.g., *I sometimes find it difficult to see things from the "other guy's" point of view; I try to look at everybody's side of a disagreement before I make a decision (R); I sometimes try to understand my friends better by imagining how things look from their perspective (R)*), and (2) the *Empathic Concern* scale (EC) contains seven indicators that inquire about emotional reactions to the negative experiences of others. EC taps participants" other-oriented feelings of warmth, sympathy, and concern for unfortunate others (e.g., *I often have tender, concerned feelings for people less fortunate than me (R); Sometimes I don't feel sorry for other people when they are having problems; I am often quite touched by things that I see happen (R)*) (Davis, 1983). Items are scored on a five-point Likert scale ranging from 1 = fully agree, to 5 = fully disagree. Before importing the data in M*plus* we used reverse coding on the positively worded indicators. High scores on the indicators imply high levels of empathic concern with an alpha reliability of.73 and perspective-taking with an alpha reliability of.68.

Perceived peer disapproval is measured with the following question: *"Regardless of what you have answered so far. Suppose you keep the 50 euro for yourself, how would your best friend(s) react to that?"*. A latent variable is created, consisting of two indicators. Response categories ranged from 1 (admire me for it) to 7 (criticize me for it). High scores on this latent variable imply a high level of perceived disapproval with an alpha reliability of.82.

Fear sensitivity is measured with the following question: *"Suppose you kept the 50 euro to yourself, how afraid would*

you be of getting caught?". A latent variable was created, consisting of two indicators. The response categories ranged from 1 (very afraid) to 7 (not afraid at all). We used reverse coding on the indicators so that high scores imply higher levels of fear sensitivity with an alpha reliability of.75.

Analytic Strategy

This study relies upon SEM as the primary analytic strategy. SEMs combine path analysis and factor analysis to allow the modeling of multiple dependent variables, estimation of direct and indirect effects measurement of latent factors, and estimation of measurement errors (Kline, 2016). Categorical WLSMV is chosen as the estimation method in M*plus* (Muthén & Muthén, 2012). Firstly the direct paths were tested, and secondly the indirect paths. The mediational analysis allows a better understanding of the sequence of effects that lead to certain outcomes (Kenny, 2008). The inclusion of mediators leads to a collection of direct and specific indirect effects. The bias-corrected bootstrap confidence interval method was used to examine the significance of the mediation effects in the SEM model as suggested by MacKinnon et al. (2004).

Results

Firstly, we investigated the bivariate interrelations of the theoretical constructs (supplementary materials). All constructs are significantly correlated in the expected direction.

Indirect effects
Perceived peer disapproval → anticipated guilt → intentions to steal = −0.052*** (95% CI: −0.072; −0.033)
Perceived peer disapproval → self-serving justifications → intentions to steal = −0.265*** (95% CI: −0.311; −0.219)
Empathic concern → anticipated guilt → intentions to steal = −0.025*** (95% CI: −0.039; −0.012)
Empathic concern → self-serving justifications → intentions to steal = 0.021 (ns) (95% CI: −0.022; 0.064)
Empathic perspective taking → anticipated guilt → intentions to steal = −0.047*** (95% CI: −0.066; −0.027)
Empathic perspective taking → self-serving justifications → intentions to steal = −0.167*** (95% CI: −0.214; −0.121)
Fear sensitivity → anticipated guilt → intentions to steal = −0.122*** (95% CI: −0.164; −0.080)
Fear sensitivity → self-serving justifications → intentions to steal = −0.148*** (95% CI: −0.195; −0.100)

Note: ***$p < .001$. ns = not significant – Coefficients reported in the path model are standardized (STDYX)—Estimator: WLSMV.

The results of SEM predicting intentions to steal, anticipated guilt, and self-serving justifications (direct effects) are presented in Figure 2. Firstly, the overall results show that higher levels of self-serving justifications heighten the likelihood of stealing ($\beta = .657$; $p = .000$) and that higher levels of anticipated guilt reduce it ($\beta = −.227$; $p = .000$). In addition, anticipated guilt and self-serving justifications are moderately negatively associated (*r*
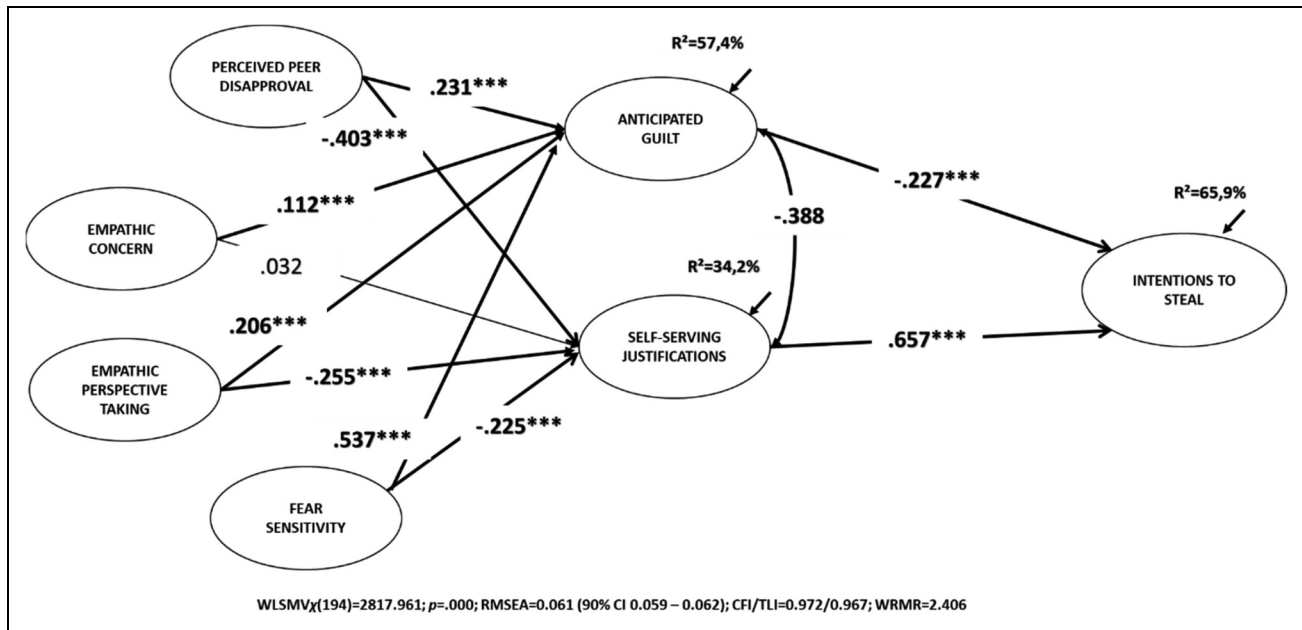
Figure 2. Results of SEM predicting intentions to steal, anticipated guilt, and self-serving justifications for the full sample (N = 3694).

= −.388; *p* = .000), the more reported self-serving justifications, the weaker reported anticipated guilt. Secondly, higher levels of self-serving justifications are significantly predicted by lower levels of perceived peer disapproval (β = −.403; *p* = .000), lower levels of empathic perspective-taking (β = −.255; *p* = .000), and lower levels of fear sensitivity (β = −.225; *p* = .000). There is no significant effect coming from empathic concern. Higher levels of anticipated guilt are significantly predicted by higher levels of fear sensitivity (β = .537; *p* = .000), higher levels of perceived peer disapproval (β = .231; *p* = .000), higher levels of empathic perspective-taking (β = .206; *p* = .000) and empathic concern (β = .112; *p* = .000). Thirdly, the results of the mediation path analysis show that all indirect effects are significant (except for the path from empathic concern to intentions to steal via self-serving justifications), in particular, the indirect paths between all independent variables and intentions to steal via self-serving justifications appear stronger than the indirect paths via anticipated guilt. Notwithstanding the differences in the size of the indirect effects, the statistical significance lends support to the hypothesis that the independent variables influence intentions to steal through anticipated guilt and self-serving justifications. Higher levels of empathic perspective-taking, perceived peer disapproval, and fear sensitivity reduce reported self-serving justifications but heighten levels of anticipated guilt, which in turn, respectively promote and curb the likelihood of stealing. Fourthly, this set of factors in the model accounts for 65.9% of the variance of intentions to steal, 57.4% of the variance in anticipated guilt, and 34.2% of the variance in self-serving justifications.

Discussion

How and why people make choices that help others, avoid hurting them, at a personal cost, have been a scientific challenge

since at least the time of Darwin (Henrich & Muthukrishna, 2021). Important mechanisms that have been put forward in explaining why individuals cooperate successfully are relatedness (Hamilton, 1964) and reciprocity (Trivers, 2006). But these do not account for why humans cooperate outside of the family group on an unparalleled scale (Apicella & Silk, 2019). Building a cooperative sociality involves suppressing the conflicts of interest between the individual group members (Raihani, 2021). The fact that individual and collective interests diverge, creates adaptive problems or fundamental social dilemmas (Dawes, 1980). Although all members of society thrive by cooperating, each individual is faced with numerous opportunities to cheat. One possible solution to solve such dilemmas and to get people to cooperate appeals to morality as credible commitment devices which allow people to gain the benefits of cooperation or protection from cheating (Dawes, 1980; Frank, 1988). The present study examined the intermediary role of anticipated guilt and self-serving justifications as facets of morality, in the relations between individual differences in empathic concern, perspective-taking, perceived peer disapproval, fear sensitivity, and intentions to steal. We find that individual differences in the aforementioned antecedents significantly and negatively predict self-serving justifications that, in turn, positively feed into intentions to steal. Conversely, we find that individual differences in empathic concern, perspective taking, perceived peer disapproval, and fear sensitivity significantly and positively predict anticipated guilt, which in turn decreases the likelihood of stealing. The latter is consistent with the idea that the experience of guilt feelings serves as a commitment device that may support individuals" propensity to cooperate (Frank, 1988). Further, we find support for the proposition that both self-serving justifications and anticipated guilt mediate the relationships between the

independent variables and intentions to steal. Noteworthy, we observe that the indirect paths via self-serving justifications have a stronger impact on intentions to steal than the indirect paths via anticipated guilt. Of course, participants" anticipated guilt feelings about imagining stealing may differ in intensity according to whom they imagine stealing the money from. If we accept the notion that guilt motivates people to treat their cared for partners well and to avoid harming them (Baumeister et al., 1994), the smaller impact via anticipated guilt might be because participants consider stealing from a stranger as an event where less weight is given to the care for that of a stranger (compared to the social value given to close kin or friends). For example, Oda and Sawada (2021) examined the relationship between anticipated guilt and moral norm violation contingent on the social relationship between the transgressor and the witness of the norm violation. Evidence was found for the proposition that the intensity of anticipated guilt depended on the nature of the social relationship. Results suggested that participants" anticipated guilt was significantly lower when the transgression was witnessed by a stranger (compared to parents, friends, or neighbors). The same line of reasoning could be extended to the relationship between the transgressor and the victim. Results might differ from what we found in our study with different types of victims, such as a close relative, a friend, or an acquaintance. In addition, the centrality of moral emotions to social adaptive problems also centers on the distinction between in-group versus out-group. Moral prohibitions such as "thou shalt not steal" often refer to members of one's in-group (Buss, 2019). This suggests that lower levels of anticipated guilt are expected at the prospect of stealing from a member of an out-group. We also observed that self-serving justifications are inversely associated with anticipated guilt. Indeed, Bandura and colleagues (Bandura et al., 1996) have long predicted that high moral disengagement (justifications) would be accompanied by low guilt and thus weakening self-restraint against antisocial behavior. Our findings are consistent with other research, although some of these studies modeled anticipated guilt as a mediator between moral disengagement and (likelihood) of antisocial behavior (Boardley et al., 2017; Kavussanu & Ring, 2017; Ring & Kavussanu, 2018; Stanger et al., 2013). This notion of morality is consistent with the argument that even people who think of themselves as good people engage in wrongdoing through various implicit self-serving processes or who use various justifications in the situation to behave badly without feeling immoral (Bazerman & Tenbrunsel, 2011), an argument that also resonates with the idea of a modular mind (Kurzban, 2011).

Study Limitations

Although the current study includes rich data from a large general sample of young people between 12 and 25 years of age and used SEM as a powerful statistical analysis technique, there are limitations. For one, our study relied on a convenience sample and self-report measures. Therefore the generalizability of the findings to other samples, using other measures, is an open question. Another important limitation relates to the use of SEMs and is termed *the ambiguity of the data* (Cliff, 1983, p. 118). The foregoing asserts that data can never positively confirm a conceptual model (Popper, 1959), they only fail to disconfirm it. There may be other plausible paths that cannot be ruled out. For instance, we modeled the paths in a recursive model, which is a straightforward model in which causal effects are unidirectional (a feature that simplifies statistical demands for the analysis) (Kline, 2016). We did not include feedback loops, such as a path from anticipated guilt to empathy. Our argument for the path between empathy (measured as a "trait" variable) and anticipated guilt (measured as a "state" variable) is consistent with previous theoretical and empirical accounts (empathy-based guilt that refrains an individual from rule-breaking behavior). However, the reverse path might be possible. For example, Leith and Baumeister (1998) found that participants" empathic response was a crucial mediator between guilt (measured as a personality trait) and actual felt guilt. Indeed, the correlational nature of our data means that caution is warranted in interpreting the findings. In addition, using SEMs in a cross-sectional research design, researchers test whether a proposed causal structure is supported by the data, but there is nothing in SEMs that magically transforms correlational data into causal conclusions (Hox & Bechger, 1998). The use of hypothetical scenarios to elicit emotions and behavioral tendencies may object that these have little relevance to real-life social situations. Fictional scenarios, in combination with questionnaires, are increasingly used in (evolutionary informed) social sciences research. Research involving hypothetical scenarios and self-report measures is most instructive in combination with other methods, like social dilemma game theory or real-life field studies. Nevertheless, the use of fictional scenarios displaying imagined situational opportunities in the context of norm-breaking events may serve to reveal the same underlying proximate psychological mechanisms that are triggered by a real-life social situation (for an extensive discussion see Wilson & O'Gorman, 2003). Although we found anticipated guilt and self-serving justifications to be strong predictors of uncooperative tendencies we certainly do not invoke them as a comprehensive answer to commitment problems. Another important factor that may influence people's uncooperative tendencies is the ability to exercise self-control. Commitment problems, such as opportunities to steal, are tempting at the moment because they offer a monetary reward as a short-term benefit but may be harmful in the long run to one's reputation, to valued relationships (Alexander, 1987; Hardy & van Vugt, 2006; Nowak & Sigmund, 2005). Besides guilt (the focus of the present study), another human emotion that makes for a deeply felt sense of right and wrong is moral shame (Boehm, 2012). In attributional theories, both moral guilt and shame give rise to the experience of painful feelings about past misdeeds but differ in scope and focus (Lewis, 1971). Whereas guilt arises from the experience of negative feelings that result from the violation of a moral code, shame results from failing to fulfill personal and social standards, and shame engulfs the whole self (Katchadourian, 2010).

Guilt has an inward focus associated with private feelings of remorse, and shame has a more outward focus, which has to do with reputational concerns because a past misdeed is known to others or may become public (Jacquet, 2016). In the evolutionary theory of moral origins put forward by Boehm (2012), shame is considered a key universal concept in the evolutionary basis of the human conscience. Especially for universalist cultures, like countries in Western Europe and America, both guilt and shame matter, but shame is more salient for people raised in collectivist cultures, like countries in Asia. In many world languages, including those of hunter-gatherers and tribal people, no moral word similar to guilt is to be found, however, shame words appear everywhere. Furthermore, shame feelings have a universal human physiological response—blushing (Boehm, 2012). An alternative, evolutionary-psychological perspective on shame stresses the adaptive function of shame. Sznycer and colleagues (Sznycer et al., 2016) propose that shame is an evolved emotion system that serves the adaptive function of defending the individual against the social devaluation of others. The prospect of being negatively viewed by others might elicit shame which neurocognitive architecture is designed to deter behaviors that would lead to more social devaluation than benefits the behavior would otherwise yield. Perceptions of peers" disapproval of an individual's intentions to steal (thus information that would cause peers to lower their social valuation of the individual) might activate shame (and other emotions that often co-occur with shame such as fear (e.g., Tangney & Dearing, 2003)). Shame and guilt can be felt together, although shame is a more global feeling than guilt (Sznycer, 2019). Given individual differences in how easily people feel ashamed (Sznycer et al., 2012), future research might further investigate whether anticipated guilt and self-serving justifications are contingent on the extent to which an individual is prone to anticipate shame towards relatives, friends, or strangers.

Conclusion

The present study presented a testable conceptual model in the explanation of intentions to steal as a moral commitment problem. We focused on the influences of anticipated guilt and self-serving justifications as facets of morality that respectively inhibit and promote participants" reported intentions to steal. We found that anticipated guilt and self-serving justifications are mediators for the influences of empathic concern and perspective-taking, fear sensitivity, and perceptions of peer disapproval on intentions to steal. Future research might apply the conceptual model to other forms of moral wrongdoings, including other types of victims (relatives, friends), or might include other moral emotions such as shame, that are relevant in solving the commitment problem.

ORCID iD

Ann De Buck https://orcid.org/0000-0002-0579-1011

Supplemental material

Supplemental material for this article is available online.

Notes

[1]Retrieved from http://www.shopliftingprevention.org (August 2022)

1. When *evolutionary biologists* speak of selfishness they mean *genetic selfishness*, whereas when *evolutionary psychologists* speak of selfishness, they mean *psychological selfishness*. The former means *fostering the propagation of the genes into the next generation,* and concerns the genetic consequences of a behavior. The latter means exclusive or excessive concern for one's own needs. The distinction between genetic and psychological selfishness is built around the principle that all that natural selection can work with are the effects of behavior, not the motivation behind it (de Waal, 2008, p. 280). Genes are little strings of DNA, devoid of motives. Genes "do what they do" without intentions or goals in mind, which means that genes cannot be selfish or unselfish (de Waal, 2022). According to Krebs (2011), psychological selfishness differs from genetic selfishness in three significant ways: (1) genetic selfishness is defined in terms of the consequences of behaviors. Psychological selfishness is defined in terms of *motivations,* a psychological state of concern, consideration for oneself, and a deficiency in concern for others; (2) when individuals behave in psychologically selfish ways, they seek to obtain an advantage, profit, and well-being for themselves. When individuals behave in genetically selfish ways, all that matters is whether that behavior contributes to the propagation of their genes, irrespective of their motivations, and (3) the standard for genetic selfishness is the relative number of genes that one contributes to future generations. The standard for psychological selfishness is an exclusive or excessive concern for one's own needs, desires as well as disregard for the well-being of others (p. 35). Sober and Wilson (1998) pointed out that there is no necessary connection between genetic and psychological forms of selfishness and altruism. Individuals who are motivated to obtain benefits for themselves without regard for others, may or may not contribute more copies of their genes to future generations relative to those who behave in less psychologically selfish ways. Nevertheless, it is safe to assume that individuals who are motivated to cooperate and to benefit themselves and others could be more likely to survive and reproduce relative to those who are motivated to seek benefits for themselves at the expense of others.

2. We stress psychological self-interest because the concept of self-interest has created a lot of confusion among scholars in different fields. Evolutionary biologists tend to genetic self-interest, which

may lead to evolutionary success. Genetic self-interest may lead to both psychological egoism and psychological altruism (see Dawkins, 2006; Sober & Wilson, 1998)

3. The visual scenarios come from the Visualising Crime Project, developed by NSCR's CRIME Lab (The Netherlands) (see: https://nscr.nl/visualizing-crime-project/). Courtesy of Prof. Dr. Jean-Louis van Gelder, director of the Department of Criminology at the Max Planck Institute for the Study of Crime, Security, and Law, Freiburg, Germany.

4. €50 is approx. 5.094 USD, €200 is approx.. 203.75 3USD (August 2022).

5. The psychometric properties of the Dutch version of the *Interpersonal Reactivity Index* (IRI) were examined by De Carte et al. (2007). The results supported the psychometric adequacy of the scores in terms of factor structure and scale reliability, construct validity as reflected in scale inter-correlations and gender differences, and discriminant and convergent validity as evidenced by correlations with other related measures although no other existing empathy measures were taken into account. In short, the study gave evidence for the reliability and validity of the Dutch version of the IRI.

References

Agnew, R. (1994). The techniques of neutralization and violence. *Criminology; An Interdisciplinary Journal*, *32*(4), 555–580. https://doi.org/10.1111/j.1745-9125.1994.tb01165.x

Alexander, R. D. (1979). *Darwinism and human affairs*. University of Washington Press.

Alexander, R. D. (1985). A biological interpretation of moral systems. *Zygon® Journal of Religion and Science*, *20*(1), 3–20. https://doi.org/10.1111/j.1467-9744.1985.tb00574.x

Alexander, R. D. (1987). *The biology of moral systems*. Transaction Publishers.

Apicella, C. L., & Silk, J. B. (2019). The evolution of human cooperation. *Current Biology*, *29*(11), R447–R450. https://doi.org/10.1016/j.cub.2019.03.036

Baker, E., Baibazarova, E., Ktistaki, G., Shelton, K. H., & Van Goozen, S. H. (2012). Development of fear and guilt in young children: Stability over time and relations with psychopathology. *Development and Psychopathology*, *24*(3), 833–845. https://doi.org/10.1017/S0954579412000399

Bandura, A. (1991). Social cognitive theory of moral thought and action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development: Theory, research, and applications* (pp. 71–129). Lawrence Erlbaum.

Bandura, A. (2017). Moral disengagement in the perpetration of inhumanities. In H. Stewart & A. Lukas Scott (Eds.), *Recent developments in criminological theory. Toward disciplinary diversity and theoretical integration* (pp. 135–152). Routledge.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, *71*(2), 364–374. https://doi.org/10.1037/0022-3514.71.2.364

Barclay, P. (2016). Reputation. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (2nd ed., Vol. 2: integrations, pp. 810–828). Wiley.

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

Baumeister, R. F. (2011). Need-to-belong theory. In P. A. M. Van Lange, Arie W. Kruglanski, & E. Tory Higgins (Eds.), *The handbook of theories of social psychology* (Vol. 2, pp. 121–140). Sage Publications Inc.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin*, *115*(2), 243. https://doi.org/10.1037/0033-2909.115.2.243

Baumeister, R. F., Vohs, K. D., DeWall, N., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review*, *11*(2), 167–203. https://doi.org/10.1177%2F1088868307301033

Bazerman, M. H., & Tenbrunsel, A. E. (2011). *Blind spots. Why we fail to do what's right and what to do about it*. Princeton University Press. https://doi.org/10.1515/9781400837991

Bloom, P. (2013). *Just babies: The origins of good and evil*. Broadway Books.

Boardley, I. D., Smith, A. L., Mills, J. P., Grix, J., & Wynne, C. (2017). Empathic and self-regulatory processes governing doping behavior. *Frontiers in Psychology*, *8*, 1495. https://doi.org/10.3389/fpsyg.2017.01495

Boehm, C. (2012). *Moral origins: The evolution of virtue, altruism, and shame*. Basic Books.

Boyd, R. (2006). The puzzle of human sociality. *Science (New York, N.Y.)*, *314*(5805), 1555–1556. https://doi.org/10.1126/science.1136841

Buss, D. M. (2019). *Evolutionary psychology. The new science of the mind* (6th ed.). Routledge.

Buss, D. M., & Shackelford, T. K. (1997). Human aggression in evolutionary psychological perspective. *Clinical Psychology Review*, *17(6)*, 605–619. https://doi.org/10.1016/S0272-7358(97)00037-8

Campbell, D. T. (1982). Legal and primary-group social controls. *Journal of Social and Biological Structures*, *5*(4), 431–438.

Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual Review of Psychology*, *66*(1), 295–319. https://doi.org/10.1146/annurev-psych-010814-015156

Čermák, I., & Blatný, M. (1995). Personality indicators of aggression and moral disengagement. *Studia Psychologica*, *37*(3), 199–201.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, *18*(1), 115–126. https://doi.org/10.1207/s15327906mbr1801_7

Cohen, L. E., & Machalek, R. (1988). A general theory of expropriative crime: An evolutionary ecological approach. *American Journal of Sociology*, *94*(3), 465–501. https://doi.org/10.1086/229027

Copes, H. (2003). Societal attachments, offending frequency, and techniques of neutralization. *Deviant Behavior*, *24*(2), 101–127. https://doi.org/10.1080/01639620390117200

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). Oxford University Press.

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 91–115). Guilford.

Cosmides, L., & Tooby, J. (2008). Can a general deontic logic capture the facts of human moral reasoning? How the mind interprets social exchange rules and detects cheaters. In W. Sinnott-Armstrong (Ed.), *Moral psychology. The evolution of morality: Adaptations and innateness* (Vol. 1, pp. 53–119). The MIT Press.

Darwin, C. (1981 [1871]). *The Descent of Man, and Selection in Relation to Sex, Vols. I & II*. Princeton University Press.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

Davis, M. H. (1994). *Empathy: A Social Psychological Approach*. Westview Press.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*(1), 169–193. https://doi.org/10.1146/annurev.ps.31.020180.001125

Dawkins, R. (2006). *The selfish gene (30th Anniversary edition)*. Oxford University Press.

De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007). Measuring empathic tendencies: Reliability and validity of the Dutch version of the interpersonal reactivity index. *Psychologica Belgica*, *47*(4), 235–260. https://doi.org/10.5334/pb-47-4-235

Detert, J. R., Treviño, L. K., & Sweitzer, V. L. (2008). Moral disengagement in ethical decision making: A study of antecedents and outcomes. *Journal of Applied Psychology*, *93*(2), 374–391. https://doi.org/10.1037/0021-9010.93.2.374

De Waal, F. B. M. (2008). Putting the altruism back in altruism. *Annual Review of Psychology*, *59*, 279–300. https://doi.org/10.1146/annurev.psych.59.103006.093625

De Waal, F. B. M. (2022). *Different. Gender through the eyes of a primatologist*. W. W. Norton & Company.

Duntley, J. D., & Shackelford, T. K. (2008). Darwinian foundations of crime and law. *Aggression and Violent Behavior*, *13*(5), 373–382. https://doi.org/10.1016/j.avb.2008.06.002

Durrant, R., & Ward, T. (2015). *Evolutionary criminology: Towards a comprehensive explanation of crime*. Academic Press.

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology*, *51*(1), 665–697. https://doi.org/10.1146/annurev.psych.51.1.665

Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 45–60). Wiley.

Frank, R. H. (1988). *Passions Within Reason: The Strategic Role of the Emotions*. W.W. Norton & Company, Inc.

Frank, R. H. (2001). Cooperation through emotional commitment. In R. M. Nesse (Ed.), *Evolution and the capacity for commitment* (pp. 57–76). Russell Sage Foundation.

Gintis, H. (2003). The hitchhiker's guide to altruism: Gene-culture coevolution, and the internalization of norms. *Journal of Theoretical Biology*, *220*(4), 407–418. https://doi.org/10.1006/jtbi.2003.3104

Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 852–870). Oxford University Press.

Hamilton, W. D. (1964). The genetical evolution of social behavior. II. *Journal of Theoretical Biology 7*(1), 17–52. https://doi.org/10.1016/0022-5193(64)90039-6

Hardin, G. R. (1968). The tragedy of the commons. *Science (New York, N.Y.)*, *162*(3859): 1243–1248. https://doi.org/10.1126/science.162.3859.1243

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%2F0146167206291006

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*, 207–240. https://doi.org/10.1146/annurev-psych-081920-042106

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

Hox, J. J., & Bechger, T. M. (1998). An introduction to structural equation modeling. *Family Science Review*, *11*(4), 354–373.

Jacquet, J. (2016). *Is shame necessary?: New uses for an old tool*. Vintage.

Kanazawa, S. (2008). Theft. In J. D. Duntley & T. K. Shackelford (Eds.), *Evolutionary forensic psychology. Darwinian foundations of crime and law* (pp. 160–175). Oxford University Press.

Katchadourian, H. (2010). *Guilt: The bite of conscience*. Stanford University Press.

Kavussanu, M., & Ring, C. (2017). Moral identity predicts doping likelihood via moral disengagement and anticipated guilt. *Journal of Sport & Exercise Psychology*, *39*(4), 293. https://www.proquest.com/scholarly-journals/moral-identity-predicts-doping-likelihood-via/docview/1988040998/se-2?accountid=11077

Kenny, D. A. (2008). Reflections on mediation. *Organizational Research Methods*, *11*(2), 353–358. https://doi.org/10.1177%2F1094428107308978

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Kochanska, G., Gross, J. N., Lin, M. H., & Nichols, K. E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development*, *73*(2), 461–482. https://doi.org/10.1111/1467-8624.00418

Krebs, D. L. (2008). Morality: An evolutionary account. *Perspectives on Psychological Science*, *3*(3), 149–172. https://doi.org/10.1111%2Fj.1745-6924.2008.00072.x

Krebs, D. L. (2011). *The origins of morality: An evolutionary account*. Oxford University Press.

Kurzban, R. (2011). *Why everyone (else) is a hypocrite*. Princeton University Press.

Leith, K. P., & Baumeister, R. F. (1998). Empathy, shame, guilt, and narratives of interpersonal conflicts: Guilt-prone people are better at perspective taking. *Journal of Personality*, *66*(1), 1–37. https://doi.org/10.1111/1467-6494.00001

Lewis, D. M., Al-Shawaf, L., Russell, E. M., & Buss, D. M. (2015). Friends and happiness: An evolutionary perspective on friendship. In M. Demir (Ed.), *Friendship and happiness. Across the life-span and cultures* (pp. 37–57). Springer.

Lewis, H. B. (1971). *Shame and guilt in neurosis*. International Universities Press.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and

resampling methods. *Multivariate Behavioral Research*, *39*(1), 99–128. https://doi.org/10.1207/s15327906mbr3901_4

Maibom, H. L. (2014). *Introduction. (almost) everything you ever wanted to know about empathy*. In H. L. Maibom (Ed.), *Empathy and morality* (pp. 9–51). Oxford University Press.

Marschall, D., Sanftner, J., & Tangney, J. P. (1994). *The State Shame and Guilt Scale*. George Mason University.

Martinez, A. G., Stuewig, J., & Tangney, J. P. (2014). Can perspective-taking reduce crime? Examining a pathway through empathic concern and guilt-proneness. *Personality and Social Psychology Bulletin 40*(12), 1659–1667. https://doi.org/10.1177/0146167214554915

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, *5*(2), 119–124. https://doi.org/10.1177%2F1754073912468165

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide: Statistical analysis with latent variables (Version 7)*. Author.

Neuberg, S. L., & DeScioli, P. (2016). Prejudices: Managing perceived threats to group life. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (2nd ed., Vol. 2: integrations, pp. 704–721). Wiley.

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*(7063), 1291–1298. https://doi.org/10.1038/nature04131

Oda, R., & Sawada, K. (2021). Do social relationships with those who witness moral transgression affect the sense of guilt? *Evolutionary Psychology*, *19*(3), 1–7. https://doi.org/10.1177%2F14747049211032576

Olthof, T. (2012). Anticipated feelings of guilt and shame as predictors of early adolescents' antisocial and prosocial interpersonal behaviour. *European Journal of Developmental Psychology*, *9*(3), 371–388. https://doi.org/10.1080/17405629.2012.680300

Panksepp, J., & Watt, D. (2011). What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emotion Review*, *3*(4), 387–396. https://doi.org/10.1177%2F1754073911410741

Pauwels, L. J., & Svensson, R. (2015). Schools and child antisocial behavior: In search for mediator effects of school-level disadvantage. *SAGE Open*, *5*(2). https://doi.org/10.1177%2F2158244015592936

Popper, K. R. (1959). *The logic of scientific discovery*. Hutchinson.

Raihani, N. (2021). *The social instinct. How cooperation shaped the world*. Vintage.

Raine, A. (1997). *The psychopathology of crime: Criminal behavior as a clinical disorder*. Academic Press.

Rawls, J. (1999). *A theory of justice: Revised edition*. Harvard University Press.

Rebellon, C. J., Manasse, M. E., Agnew, R., Van Gundy, K. T., & Cohn, E. S. (2016). The relationship between gender and delinquency: Assessing the mediating role of anticipated guilt. *Journal of Criminal Justice*, *44*, 77–88. https://doi.org/10.1016/j.jcrimjus.2015.11.006

Ribeaud, D., & Eisner, M. (2010). Are moral disengagement, neutralization techniques, and self-serving cognitive distortions the same? Developing a unified scale of moral neutralization of aggression. *International Journal of Conflict and Violence (IJCV)*, *4*(2), 298–315. https://doi.org/10.4119/ijcv-2833

Richerson, P., & Boyd, R. (1998). The evolution of ultrasociality. In I. Eibl-Eibesfeldt & F. K. Salter (Eds.), *Indoctrinability, ideology, and warfare* (pp. 71–96). Berghahn Books.

Ring, C., & Kavussanu, M. (2018). The role of self-regulatory efficacy, moral disengagement, and guilt on doping likelihood: A social cognitive theory perspective. *Journal of Sports Sciences*, *36*(5), 578–584. https://doi.org/10.1080/02640414.2017.1324206

Sapolsky, R. M. (2017). *Behave. The biology of humans at our best and worst*. Penguin.

Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science*, *24*(2), 125–130. https://doi.org/10.1177%2F0963721414553264

Siebert, E. C., & Stewart, D. G. (2019). Neutralization technique use predicts delinquency and substance use outcomes. *Journal of Substance Abuse Treatment*, *102*, 8–15. https://doi.org/10.1016/j.jsat.2019.04.006

Smith, A. (1817 [1759]). *The Theory of Moral Sentiments, Vols. I & II*. Wells and Lilly.

Sober, E., & Wilson, D. S. (1998). *Unto others*. Harvard University Press.

Somerville, L. H. (2013). The teenage brain: Sensitivity to social evaluation. *Current Directions in Psychological Science*, *22*(2), 121–127. https://doi.org/10.1177%2F0963721413476512

Stanger, N., Kavussanu, M., Boardley, I. D., & Ring, C. (2013). The influence of moral disengagement and negative emotion on antisocial sport behavior. *Sport, Exercise, and Performance Psychology*, *2*(2), 117–129. https://doi.org/10.1037/a0030585

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, *28*(1), 78–106. https://doi.org/10.1016/j.dr.2007.08.002

Sykes, G., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review*, *22*(6), 664–670. https://doi.org/10.2307/2089195

Sznycer, D. (2019). Forms and functions of the self-conscious emotions. *Trends in Cognitive Sciences*, *23*(2), 143–157. https://doi.org/10.1016/j.tics.2018.11.007

Sznycer, D., Takemura, K., Delton, A. W., Sato, K., Robertson, T., Cosmides, L., & Tooby, J. (2012). Cross-cultural differences and similarities in proneness to shame: An adaptationist and ecological approach. *Evolutionary Psychology*, *10*(2), 147470491201000213. https://doi.org/10.1177%2F147470491201000213

Sznycer, D., Tooby, J., Cosmides, L., Porat, R., Shalvi, S., & Halperin, E. (2016). Shame closely tracks the threat of devaluation by others, even across cultures. *Proceedings of the National Academy of Sciences*, *113*(10), 2625–2630. https://doi.org/10.1073/pnas.1514699113

Tangney, J. P. (1990). Assessing individual differences in proneness to shame and guilt: Development of the self-conscious affect and attribution inventory. *Journal of Personality and Social Psychology*, *59*(1), 102–111. https://doi.org/10.1037/0022-3514.59.1.102

Tangney, J. P., & Dearing, R. L. (2003). *Shame and guilt*. Guilford Press.

Tangney, J. P., & Stuewig, J. (2004). A moral-emotional perspective on evil persons and evil deeds. In A. G. Miller (Ed.), *The social psychology of good and evil* (pp. 327–355). The Guilford Press.

Tomasello, M. (2016). *A natural history of human morality*. Harvard University Press.

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, *64*(1), 231–255. https://doi.org/10.1146/annurev-psych-113011-143812

Tooby, J., & Cosmides, L. (1996). Friendship and the banker's paradox: Other pathways to the evolution of adaptations for altruism. In W. G. Runciman, J. M. Smith, & R. I. M. Dunbar (Eds.), *Evolution of social behavior patterns in primates and man* (Vol. 88, pp. 119–143). Oxford University Press.

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 114–137). Guilford.

Trivers, R. (2002). *Natural selection and social theory: Selected papers of Robert Trivers*. Oxford University Press.

Trivers, R. (2006). Reciprocal altruism: 30 years later. In P. M. Kapeller & C. P. Van Schaik (Eds.), *Cooperation in primates and humans: Mechanisms and evolution* (pp. 67–83). Springer. https://doi.org/10.1007/3-540-28277-7_4

Turchin, P. (2015). *Ultrasociety: how 10000 years of war made humans the greatest cooperators on earth*. Beresta Books.

Van Gelder, J. L., & De Vries, R. E. (2014). Rational misbehavior? Evaluating an integrated dual-process model of criminal decision making. *Journal of Quantitative Criminology*, *30*(1), 1–27. https://doi.org/10.1007/s10940-012-9192-8

Walsh, A. (2010). *Biology and criminology: The biosocial synthesis*. Routledge.

Walsh, A. (2018). *Reinforcement sensitivity theory: a metatheory for biosocial criminology*. Routledge.

Walsh, A., & Jorgensen, C. (2018). Evolutionary theory and criminology. In R. Hopcroft (Ed.), *Oxford Handbook of evolution, biology, and society* (pp. 517–542). Oxford University Press.

Walters, G. D. (2020). Crime and social cognition: A meta-analytic review of the developmental roots of adult criminal thinking. *Journal of Experimental Criminology*, *18*, 1–25. https://doi.org/10.1007/s11292-020-09435-w

Walters, G. D., Runell, L., & Kremser, J. (2021). Drifters and Delinquents: Commitment and Neutralization in Low- and High-Rate Offending Early Adolescents. *Crime & Delinquency*. https://doi.org/10.1177/00111287211039992

Wilson, D. S., & O'Gorman, R. (2003). Emotions and actions associated with norm-breaking events. *Human Nature*, *14*(3), 277–304. https://doi.org/10.1007/s12110-003-1007-z