

Implicit Theories of Morality, Personality, and Contextual Factors in Moral Appraisal¹

Ana-Maria Hojbotă

Abstract: This article explores the implicit theories of morality, or the conceptions regarding the patterns of stability, continuity and change in moral dispositions, both in lay and academic discourses. The controversies surrounding these conceptions and the fragmentation of the models and perspectives in metaethics and moral psychology endangers the pursuit of adequate operationalizations of morally relevant constructs. The current debate between situationists, who deny that character is an useful concept for understanding human behavior, which is better explained by contextual factors (Doris 1998; Harman 1998) and dispositionists, who advocate the cross-situational stability of traits, is also present in the lay discourse, through the existence of competing commonsense ontological assumptions regarding the mutability or alterability of moral features, namely the implicit theories perspective (Chiu, Dweck, Tong, and Fu 1997). These personal theories are the primary suspects in the affective and cognitive reactions to transgressions: the type of attended information in formulating evaluative judgments, the calibration of moral responsibility and blameworthiness, the assignment of retribution or reparatory recommendations to transgressors. In the second part of the study we attempt to advance toward a more fine-grained inspection of these lay beliefs, arguing that the construct of implicit theories of morality, as it is currently treated and measured, tends to be restrictive and oversimplifying.

Keywords: implicit theories, folk conceptions of morality, situationism, dispositionism

Introduction

Research on the sources of variability regarding the attribution of moral traits and the views on moral character, either implicit or explicit, cannot be complete, nor entirely legitimate without situating the different perspectives into corresponding philosophical frames. Research in social psychology on person perception, more specifically, on implicit theories of morality (Chiu, Dweck, Tong, and Fu 1997; Dweck 1991; Dweck, Chiu, and Hong 1995; Miller, Burgoon, and Hall 2007), and on essentialist beliefs (Haslam, Bastian, Bain, and Kashima

¹ This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/159/1.5/133675.

2006) reveals that people generally tend to be oriented either towards a situational, contextual view of moral conduct, or a more trait-based vision of morality. This strand of research has been generally conducted without paying attention to philosophical accounts that discuss concepts such as character, human nature, traits, virtues, vices and the complicated relation between the individual and environment. This isolated approach led to a weak psychological account of moral lay thinking and to an inflation of terms and concepts from the moral domain.

Bridging the domains of personality, person perception and moral psychology, these theories are important for all three directions of research, bringing light upon the consistency of moral behavior, on how people interpret morally relevant situations and make assumptions about the people they observe and how identity builds on these types of mechanisms.

Spontaneous or analytical descriptive statements of character, explanations and predictions of future behavior are filtered by these interpretative lenses that some authors call implicit or naïve theories. They are supposed to calibrate the amount of blame, responsibility and the valence of normative judgments regarding human behaviors. Before discussing the implications of each perspective, we will first inspect the main philosophical accounts that diverge on the legitimacy and explanatory power of character traits and of situations. The fragmentation of the models describing the nature of human morality points to the inherent complexity of the phenomena. There is considerable variability in the views of the lay thinker, which points to the difficulties faced by the scholars in the field of morality. Confronted with the atomization of perspectives, models and operational definitions, the novices in this field may be discouraged when they approach the issue of moral evaluative judgments.

Based on the suggestions offered by the philosophical literature and the debates that emerged in psychology, we propose a way in which the understanding of implicit theories of morality could be improved, arguing for a more accurate measurement of moral theories and conceptions. We extend Dweck's model of implicit theories applied to morality in a few critical ways: a) by discussing the nature of moral trait attributions, as an automatic, pervasive process that inform all the rational and unconscious operations in behavior evaluation; b) by suggesting that the approach of moral traits should take into account the valence of these traits, since they activate different representational information; c) by segregating the incremental-entity dichotomy into further dimensions, relating to two sets of distinct but associated phenomena: the opposition between processes of consolidation and those related to degradation of character traits; distinguishing among morally blameworthy and praiseworthy changes in behavior, virtues and character strengths, on the one hand, and bad habits or vices on the other hand.

The Moralistic Instinct. The Pervasiveness of Moral Attribution

“The only difference between the saint and the sinner is that every saint has a past, and every sinner has a future.” Oscar Wilde’s popular quote could be understood as an indication of the human tendency towards evaluating people based on character labels. It also depicts the relativity of these traits in describing persons and predicting a future moral conduct. This kind of dichotomous thinking, however, seems to be especially appealing for the human mind. People need to believe in saints or sinners, heroes or villains, and their mind seems to be automatically attuned to identifying antagonists, be they scapegoats or saviors, tendencies that are well reflected in the vast majority of cultural products. Wilde’s words also point to the inherent complexity of assessing moral behavior and self-regulation, a feature which discouraged eminent psychologists from approaching the issue, such as the founding father of personality psychology, Gordon Allport (1937). Allport overtly expressed a view according to which morality cannot be a suitable object of psychological inquiry. Without adhering to this kind of skepticism, we approach this domain acknowledging the lack of consensus, incomplete explanatory models and paradoxical findings.

Some philosophers attempted to reject the magnetic appeal of reductionist virtue-based conceptions as those suggested by Wilde’s words. For instance, Christine Korsgaard mockingly labels the “good dog” view of the moral person in contrast to the similarly popular prototype of the “reformed miserable sinner.” According to Korsgaard, the “good dog” view of the virtuous agent as an individual equipped with “desires and inclinations” that “have been so perfectly trained that he always does what he ought to do spontaneously and with tail-wagging cheerfulness and enthusiasm” (Korsgaard 2009, 3) contrasts the second view, according to which the individual continuously faces “unruly desires in order to conform to the demands of duty” (Korsgaard 2009, 3). These two views are both theoretically sterile and methodologically inoperable and illustrate the inherent tendencies and limitations of folk epistemology, in an area where humans seem to be innately equipped with the necessary assessment tools for discerning “good dogs” from bad ones. Although rudimentary, these evaluative tools show high sensitivity to a variety of factors, prioritizing egotistical and evaluative motivations over epistemic ones.

Since the onset of the cognitivist revolution, social psychologists discuss the tendency to attach moral intentions, emotions and moral character traits, even to inanimate objects. For instance, Heider and Simmel’s experiment from 1944 asked subjects to describe geometrical shapes that executed successive movements, in an attempt to present “situations and activity without the face” (Heider and Simmel 1944, 244). A large triangle, a small triangle and a circle moved in various directions during the 2½ min film presented to the participants. Subjects were asked to describe the figures as if they were persons (e.g. “What kind of person is the big triangle?”). Although providing various

scenarios for the displayed scenes, participants judged the “characters” of geometrical shapes “with great uniformity” (Heider and Simmel 1944, 248), either as aggressive, bad-tempered, troublesome or as mean, in the case of the big triangle, or heroic and courageous, as the small triangle was frequently depicted.

This tendency to make normative (right/wrong) or axiological (good/bad) judgments is supposed to evolve very early in the course of ontogenetic development, prior to the development of language. This was illustrated by a study conducted by Hamlin, Wynn and Bloom (2007). In this experiment, toddlers of six and ten months watched a series of visual scenarios. They designed a set of simple scenes, where the main characters were “interpreted” by large wood shapes with large googly eyes, struggling to climb up a hill. Other shapes either supported and helped the climber, or trumped his climbing efforts, throwing them down. When the experiment ended, the toddlers reached for the helpers, indicating a significantly greater preference for them than for the hinderers. The authors concluded that the preference for pro-social actors is universal and innate, and manifests itself prior to the acquisition of language or higher-order reasoning. These studies indicate that people readily form impressions of the objects they observe and that these impressions are mainly translated in characterial terms:

The capacity to evaluate individuals by their social actions may also serve as a foundation for a developing system of moral cognition. Plainly, many aspects of a full-fledged moral system are beyond the preverbal infant. Yet the ability to judge differentially those who perform positive and negative social acts may form an essential basis for any system that will eventually contain more abstract concepts of right or wrong. The social evaluations we have observed in our young subjects have (at least) one crucial component of genuine moral judgments: they do not stem from infants’ won experiences with the actors involved. Our subjects had no previous history with our characters, nor did they themselves experience any consequences of these characters’ actions. Their evaluations were made on the basis of witnessed interactions between unknown individuals: the infant, as an unaffected, unrelated (and therefore unbiased) third party, is nonetheless rendering a judgment about the value of a social act (Hamlin, Wynn, and Bloom 2007, 558-559).

The authors maintain that this basic evaluative competence is based upon the fact that people constantly need to “make accurate decisions about who is friend and who is foe, who is an appropriate social partner and who is not” (Hamlin, Wynn, and Bloom 2007, 557). This tendency towards automatic moral evaluation is comprehensively discussed by Jonathan Haidt in his explorations on self-righteousness.

Jonathan Haidt is the proponent of a multidimensional notion of morality, as an evaluative “organ” with six kinds of different receptors, called moral foundations. He describes a propensity that looks like an obsession for “self-righteousness,” a tendency that, in his opinion, provided human species with the

possibility of forming cooperative groups but also led to one of the human nature's greatest flaw, the tendency of people to become hypocrites, convinced of their own virtue. This approach is best developed in his latest and controversial book, *The Righteous Mind. Why Good People Are Divided by Religion and Politics*, where he claims that human nature is not only intrinsically moral, but also highly moralistic, in other words, has a natural tendency to be critical, self-righteous, and judgmental. Thus, righteousness becomes a central feature of the human thinking, especially about social objects. Explaining its etymology, Haidt shows that the term comes from the Hebrew term *Tzedek* that refers to God's moral evaluation of people. This divine act of judgment, specific to the creator, is regarded as harsh, but also inherently just. The word applies to the way people reflect on the acts of others, including the meaning of intransigence and impartiality.

This universal tendency is also expressed by other authors, who maintain that human beings have an embedded tool of moral assessment that evolved to accurately detect and deter potential enemies, or persons that could act in ways that could be leaving intentional harm upon them. Some authors maintain that all ethical decisions can be seen as character-related: "the basic judgments in ethics are judgments about character," a view that has been gradually gaining currency (Statman 1997, 7). This tendency makes sense in an evolutionary framework.

People who cause harm intentionally are, in general, far more dangerous than people who cause harm accidentally, and therefore is more important to deter them (Greene 2013).

Intentionality thus becomes a central feature of this evaluative moral module. Joshua Greene points also to the fact that naïve moral epistemology, which is inherently dominated by this disposition towards self-righteousness, can be conceived as coherentist rather than foundationalist (Greene 2013). People rarely tend to judge the morality of an action in isolation, based on the proof regarding intentionality and blame alone, and the temptation to appeal to character judgments is overtly recognized and sometimes actively controlled. For instance, in some legal contexts, jurors are not provided evidence regarding the record of the defendant in order to avoid biased evaluation. We further discuss to what extent people endorse this dichotomous pattern and thinking about character, its adequacy and other issues related to naïve moral thinking.

Saints and sinners, situationists and dispositionists

While there is growing evidence that the morphology of folk moral psychology, at least in Western contexts, tends to be trait-based, it is not clear how lay people conceptualize the direction and depth of the changes that orient the development of moral character. It is this specific type of views on behavioral modification that vary from skeptical to more melioristic positions, that we are

expecting to affect positions regarding the allocation and punishment, and the shape of educational and remedial interventions.

In their paper “Sinning Saints and Sainly Sinners: The Paradox of Moral Self-Regulation,” Sachdeva, Iliev and Medin (2009) pointed that many moral as well as immoral behaviors stem from people’s attempts at finding the balance between competing forces, stressing that rather than labeling people as either sinners or saints, a more accurate view would be to say that people are frequently swinging between “sinning saint” and “sainly sinner” states. This inherent duality is also metaphorically expressed in an expression proposed by Phillip Zimbardo, “the Lucifer effect,” largely discussed in his homonymous book. Zimbardo seems to embrace a compromise view of the human nature, proposing a model of “situated character transformations,” that aims to push the reader further away from “the comfortable separation of Your Good and Faultless Side from Their Evil and Wicked Side” (Zimbardo 2007, 3). This account resembles more recent accounts in social psychology that sanctions absolute, isolated evaluations of people’s behavior, stemming only from universal principles, and disregarding the context. Such a model, the socio-relational model of morality (Rai and Fiske 2011), indicates that all moral practices and motives are contextualized and that behaviors can be judged either as moral or immoral depending on the perspective from which they are interpreted.

But unlike the fair to mild, quotidian or “normal” transgressions or deviations from normality described by Sachdeva, Iliev and Medin (2009), Zimbardo embarks on a more grim analysis of character, advocating what he calls three psychological truths:

First, the world is filled with both good and evil – was, is, will always be. Second, the barrier between good and evil is permeable and nebulous. And third, it is possible for angels to become devils and, perhaps more difficult to conceive, for devils to become angels (Sachdeva, Iliev, and Medin, 2009, 3).

However, although it might invite to more mystification of the issues brought under scrutiny, we read Zimbardo’s observations as pleas for achieving more clarity and cultivating a more nuanced view of the human nature than what commonsense, religion or even some scholars usually advert. For instance, the author states that this kind of reasoning in which evil is essentialized and “becomes an entity, a quality that is inherent in some people and not in others” (Zimbardo 2007, 6) may be dangerous for various reasons, such as: “upholding a Good-Evil dichotomy also takes ‘good people’ off the responsibility hook” (Zimbardo 2007, 6). Zimbardo illustrates this by resorting to Stanley Milgram’s experiment at Yale University and his own Stanford prison experiment, where researchers who recruited psychologically healthy, randomly selected subjects, showed how, when facing certain situations, people are capable of terrible things.

This idea was earlier argued and analyzed through the lens of Greek philosophy and tragedy by Martha Nussbaum in *The Fragility of Goodness*,

where she argued that “goodness of character makes the good life tolerably stable in the face of the world. But this stability is not limitless. There is a real gap between being good and living well; uncontrolled happening can step into this gap, impeding the good state of character from finding its proper fulfillment in action.” (Nussbaum 1986, 334). Nussbaum tries to argue in this book how goodness can be protected from the pull of these “uncontrolled happenings.”

Creating laboratory settings for this kind of “uncontrolled happenings” has become the systematic tool of investigation in social experimental psychology. Most of these investigations actively advocate against the idea of cross-situational consistency of moral behavior and for a sensitivity of human conduct to contextual factors that makes, on the one hand, the goal of judging the moral status of human beings futile, and on the other, the concepts of moral trait or virtue scientifically inoperant.

The tendency to see people in terms of good or bad is more accentuated in some cultures, professions or contexts and there is a long ongoing debate in philosophy regarding the legitimacy of using the notions of traits and characters. Some defenders of situationism, such as Harman (1999), propose that folk moral psychology is a character trait-based one, resembling virtue ethics. The ubiquitous presence of saint, sinners and villains in the lay discourse shows that traits are appealing not only in isolation, but also as typologies of moral characters that reflect the idea of exceptionality. Also, these scholars are overtly denouncing the plausibility of character as a valid construct with explanatory and predictive power, and consider it a reified, social construct (also see Alfano 2013; Doris 2002; Harman 1999).

Both Zimbardo and Nussbaum point to the importance of situational factors, which were, according to the situationists, ignored by generations of psychologists unilaterally concerned with measuring and ascribing static diagnoses (Alfano 2013; Doris 2002; Harman 1999). The criticism mirror the prior Anscombe’s (1958) accusations targeting the defenders of Kantianism and utilitarianism, whom he blamed for externalizing the grounds of human morality into norms and contexts. Aretaic ethics develops a dispositional and elitist view on moral excellence (in Latin: “virtus,” in Greek: “arete”). Aristotle best reflected this view in the following passage from his *Nicomachean Ethics*:

anyone can get angry – that is easy – or give or spend money; but to do this to the right person, to the right extent, at the right time, with the right aim, and in the right way, that is not for everyone, nor is it easy; that is why goodness is both rare and laudable and noble (Aristotle, NE 2.9, 1109 a26-29).

Ethical behavior is seen as resulting from displaying such virtues that, according to Aristotle, can be learned and developed into habits. Anscombe (1958) claimed that modern philosophical accounts, namely the two perspectives mentioned above, eliminated or derogated the Greek moralist’s idea of virtue, because they insist on principles, and seek to reveal the sources and contents of norms, duties,

laws and obligations, rather than to investigate those traits and habits people should cultivate in order to live a good life.

Anscombe's provocative assumptions stirred the spirits both in psychology and philosophy and generated more questions on how these issues should be addressed. At the same time with the rising popularity of experimental social psychology, concerned with finding the mediating and moderating situational factors for moral conduct, Aristotle's virtue ethics gained renewed interest in psychological research and practice with the advent of positive psychology. For instance, the most notable and used model is the one comprehensively presented in Peterson and Seligman's *Character Strengths and Virtues* (Peterson and Seligman 2004). The model reflects recent trends in personology and personality psychology, more specifically the increasing adherence to dimensional models. The categories, including strengths and virtues, are organized along a six-units dimensional model of "moral excellence" or virtues (wisdom and knowledge, courage, humanity, justice, temperance, and transcendence), based on historical surveys, and further divided into twenty four character strengths (e.g. bravery, kindness, and perseverance), which are considered to be intrinsically valued, trait-like, and endorsed by societal norms and institutions. Although virtues may be absent in some individuals, they are

perhaps grounded in biology through an evolutionary process that selected for these aspects of excellence as means of solving important tasks necessary for the survival of the species (Peterson and Seligman 2004, 13).

In the introduction to this "manual of sanities" (a label borrowed from Easterbrook 2001, 23), the authors clearly state that they "believe that good character can be cultivated, but to do so, we need conceptual and empirical tools to craft and evaluate interventions" (Peterson and Seligman 2004, 3), thus expressing their adherence to the traditional perspective introduced by Aristotle, virtue ethics.

From the situationist side, John Doris argues that good character should buffer against situational forces that threaten the occurrence of trait-relevant behaviors. His understanding of dispositionism supposes this almost universal capacity of traits to ward off unflattering reactions within "uncontrolled happenings," as Nussbaum called these situations:

When a person has a robust trait, they can be confidently expected to display trait-relevant behavior across a wide variety of trait-relevant situations, even when some or all these situations are not optimally conducive to such behavior (Doris 2002, 18).

Doris' critique of stability and globalism is based on the low consistency of the correlations, the marked impact of situational factors on behavior that "undermines the attribution of robust traits" (Doris 2002, 64) and the low correlations between personality measures and behavioral observations. In addition to this, Doris also discusses the ideographical, biographical accounts

that point to the phenomenon of trait disintegration. Besides the various criticisms that could be brought to these studies' validity of results, the argument proposed by Doris against virtue ethics does not hold, since Aristotle does not advocate an all-encompassing stability of virtue and does not deny the vulnerability of human conduct to "unobtrusive situational factors."

In a volume dedicated to defending the situationist perspective, Alfano (2013) restates the majority of this view's claims in the light of recent research. He argues that the assumptions of virtue ethics, according to which traits are stable, have trans-situational consistency, explanatory/diagnostic, and predictive potential are mostly invalidated by the experimental accounts that reveal the sensibility of human judgments and action to contextual factors, such as moods and states induced by various means, including the administration of hormones and neurotransmitters (serotonin, oxytocin), or environmental sounds and smells. The illusion of traitedness is generated, according to Alfano, by the fact that the attribution, function as self-fulfilling prophecies for the evaluated person. The mere fact of trait-labelling, enable that people's evaluation create and promote artificial traits or "factitious virtues," which are different from the virtues as they are understood by neo-Aristotelian supporters, because these artificially created virtues require social reinforcement and do not possess autonomous motivational force over their owner.

Some authors stress that such polemics surrounding virtue ethics are artificial. Kupperman (2009) gave a convincing demonstration on why situationism cannot base its claims against the existence of genuine virtues on the results of recent social psychology experiments. The appeal to studies that indicate the sensibility of moral behavior to exterior influences does not automatically function as an argument against the traits approach or against the merits of virtue ethics, but certainly motivate psychologists to gather more evidence for stability, to find other means to clarify and establish the validity of the construct. The problem is complicated by the fact that situationists tend to defend a perspective of virtues that does not necessarily reflects its founding fathers:

the character traits conceived of and debunked by situationist social psychological studies have very little to do with character as it is conceived of in traditional virtue ethics (Kamtekar and Rachana 2004, 460).

Acknowledging individual differences in moral dispositions (character traits or strengths, virtues and vices, sensitivities and weaknesses) does not say anything about their modifiability. In the light of the situationist rejection of dispositions, one can also easily disregard efforts aimed at enhancing moral behavior and strengthening of moral traits, an endeavor mostly advertised by positive psychology.

More recent integrative models of moral personality attempt to integrate the tensions between understanding the dispositional bases of moral traits, such as altruism (Carlos et al 2009). and virtues, such as gratitude (Emmons 2009).

This trend of research argues that personality can be properly and comprehensively described as

(1) an individual's unique variation on the general evolutionary design for human nature, expressed as a developing pattern of (2) dispositional traits, (3) characteristic adaptations, and (4) self-defining life narratives, complexly and differentially situated in (5) culture and social context (McAdams 2009, 12).

In other words, any approach of moral personality has to take into account three levels: dispositional traits – the heritable, longitudinally stable aspect, adaptations (that offer more subtle, contextualized explanation of behavior, integrating moral goals and schemas), and life stories, or how people conceive the good life and attribute meaning to negative life events.

The research reviewed above suggests that human beings are predisposed to implicit, automatic evaluations of actions and offer factual descriptions of their past and predictions of their future behaviors as well. Thus, the ability to evaluate the predisposition of someone to do the good or right thing and to identify someone's moral strengths and shortcomings is primary. Prior to selecting and applying moral principles to the current situation or making utilitarian judgments, people are first evaluating others and themselves in relation to some aspiring models of what they consider defining for a moral and good person. Unfortunately, there is great variability in these conceptions, variability that is also visible in the academic definitions of the issue. Psychology and philosophy still struggle to provide sound conceptions of "personality and motivational structure it expects of morally mature individuals." (Flanagan 1991, 35)

Returning to our central question, we must say we are not concerned with finding the best explanatory model for moral behavior and dispositions. We don't ask whether people are accurate in their conceptions about moral character (of course, we cannot ignore the fact that it is difficult to establish provisory accuracy in the absence of a wide agreement between scholars). Instead, we will focus on how people think of moral personality. People are intuitive virtue ethicists and, most often, the way they define prototypes of moral excellence and virtues, is self-serving. We agree with DeSteno and Valdesolo (2011), who, when discussing the malleability of character, argue that:

Character is the currency we employ to make judgments about people – to determine who is good and who is flawed, who is worthy and who is not, who is saved and who is damned (DeSteno and Valdesolo 2011, 6).

We intend to discuss the dimensionality and realism of these commonsense frames of reasoning. Our aim is to examine where lay people situate in relation to these "theoretical" dichotomies. Our main investigation tool is the construct of implicit theories. However, to do so, we need to make sure we operate with a valid and sufficiently well-defined construct. The first step in construct validation is providing consensual definitions for constructs. This

requires a parallel examination of the contents of implicit theories of moral characteristics, but also of the empirical evidence related to moral features themselves, as it may be found in mainstream psychological accounts. At this moment, trait-based models emerging in the virtue-ethics tradition provide a rather weak support for the idea of moral traits.

The Role of Character Traits and Situations in Commonsense Understandings of Morality

The idea of moral behavior as stemming from a set of dispositional features is characteristic for at least two accounts that dominate the modern psychological discourse: dispositionism and situationism. Dispositionism allows a working definition of moral characteristics, while situationism, although it does not deny their role and functionality at the lay discourse level, it denies their legitimacy in the field of scientific inquiry.

Among all the competing perspectives in moral psychology, we favor the socio-cognitive perspective of personality, which operates with cognitive-affective units or modules of mental representations that are not context-dependent, but are sensitive to certain enabling environmental stimuli. These include self-schemas, beliefs, construal of psychological and social phenomena and situations, hierarchically organized personal goals, expectations and values, self-regulatory skills and competences (Shoda, Tiernan, and Michel 2002). These models organize and provide integrative and complex accounts of trait domains, saving moral psychology from the oversimplifying trait approach and also from the tyranny of situation.

Folk moral judgments are based on a system of norms endorsed by the perceivers and on their assumptions and beliefs about people's capacities, which prove to be sophisticated and resilient (Guglielmo, Monroe, and Malle 2009). These judgments are often influenced by specific cultural, widespread beliefs regarding either theories of the world and how justice is established, or the psychological constitution of moral agents. People also differ in their assumptions regarding the fairness of the world. For instance, people who believe that "what goes around, comes around" endorse the just-world fallacy, meaning that they consider that actions are always attracting morally fitting consequences that are dictated by a superior moral force that establishes this balance (Furnham 2003). A related belief, immanent justice reasoning (for a review, see Callan, Sutton, Harvey, and Dawtry 2014) endorses a desert-based view according to which a person's fortunate or unfortunate outcomes are causally connected to that individual's prior moral behavior ("You reap what you sow"). According to this conception, a person's moral conduct, either governed by virtue or by chance, will be sooner or later sanctioned by external, immanent forces.

What individuals think of what, who and how is moral is important for at least two reasons. On the one hand, moral ideals can be extracted and inferred

from moral definitions. People hold certain prototypes of moral personality, as Lapsley and Lasky (2001) showed. Secondly, Osswald, Greitemeyer, Fischer and Frey (2010) indicate that traits are not only descriptive, inert entities, but indicated that their activation influences behavior.

As already pointed out, people have a sensitive moral compass that tends to be calibrated by the type of situation and influenced by the fact that the person is the also the agent or just observer. Studies on social perception indicate that people think in terms of traits, which manifest robustly across situations (Jones 1990; Ross and Nisbett 1991). Based on the reviewed empirical research, Doris (2002) maintains that character attributions are generally “undercontextualized and overconfident” (Doris 2002, 97).

This, however, tends to vary from a culture to another (Church, Ortiz, Katigbak, Avdeyeva, Emerson, Vargas, and Ibanez 2003; Doris 2002). Independence and interdependence or individualism and collectivism are supposed to influence the structure of these types of beliefs and claims:

The Western conception of the person as a bounded, unique or less integrated and cognitive universe, a dynamic center of awareness, emotional, judgment and action organized into a distinctive whole and set contrastively both against other such wholes and against its social and natural background is, however incorrigible it may seem to us, a rather peculiar idea in the context of the world's cultures. (Geertz 1984, 126).

Applied to the moral domain, this could be reflected as the degree to which people believe in human moral perfectibility, how flexible the capacity to overcome habits and vices is, what can be accounted as an acceptable deviation from morality and to what degree responsibility is shared.

Nevertheless, there is not only significant intercultural, but also great interpersonal variability regarding how people decide who, how and when is good or bad, what are the sources of normativity upon which their judgments rely, or how they conceive moral traits, namely their features (innateness, stability, educability, globality) and also their contents (“one person’s ‘integrity’ is another person’s ‘stubbornness,’ [and one person’s] ‘honesty’ in expressing your true feelings’ is another person’s ‘insensitivity to the feelings of others’” (Kohlberg and Mayer 1972, 479).

Meindl and Graham (2014) call this variability ‘lay moral disagreement’ and identify three forms in which disagreement manifests. The first two reflect interpersonal and intercultural variability in the way behavior and traits occur in moral judgments: the degree to which people find traits and actions as morally relevant and the possibility of attaching moral valence to actions and behavior (Graham, Haidt, Koleva, Motyl, Iyer, Wojcik, and Ditto 2013). The third lay moral disagreement is situated at the interface between naive and scientific formulations and it refers to the way traits and behaviors are invested with moral relevance.

Meindl and Graham (2014) find it important to begin the study of morality at this last level and recommend to their fellow scholars to pretest the measures that operationalize their research variables. The results presented by the two authors unfortunately indicate that lay participants do not share the conception of researchers regarding the moral prototypicality of traits or specific behaviors. The authors provide results of their own surveys indicating which attributes and acts enjoy wide-spread agreement regarding their inclusion in the area of morality (e.g. fairness), and which traits and behaviors are rather low in inter-rater consistency and thus, are problematic operationalizations of morality and need to be avoided (e.g. cooperativeness, helpfulness). Other studies show that there is also great inter-individual variability regarding meta-ethical aspects, such as the degree to which laypersons employ moral relativism in their evaluations (Gill 2009; Sarkissian, Park, Tien, Wright, and Knobe 2012).

A class of folk constructions that we take special interest in refers to the beliefs in the stable versus incremental nature of character. There are two different views: one sees human nature as constant and stable (entity theory); the other endorses a more malleable view on character (incremental theory). People who endorse entity theories tend to understand outcomes of actions in terms of traits, while incremental theorists appeal to psychological mediators, situational factors, contextual relations that may have led to the outcome.

Lay/Implicit Theories of Social Categories. The Special Case of Moral Constructs

The consistency aspect of the concept of moral disposition is notoriously disputed and complex. People's beliefs on what is moral differ in regards to how they allocate importance to personal in addition to contextual factors in interpreting the rightness of actions. Essentialist lay beliefs understand social entities (e.g. race) as rigid entities, with fixed, immutable and discrete features that are usually thought as innate, hardwired aspects with biological bases (Bastian and Haslam 2006; Williams and Eberhardt 2008). The literature on essentialist beliefs converged towards the conclusion that these types of lay theories can lead to specific, negative social categorization processes, such as stereotypes and prejudice (Levy, Stroessner, and Dweck 1998; Plaks, Stroessner, Dweck, and Sherman 2001), perpetuation of beliefs that manifest as grounds for system justification of inequalities (Keller 2005; Verkuyten 2003; Williams and Eberhardt 2008), and other related phenomena, such as dehumanization of out-group members (Leyens et al. 2000).

This rigid type of processing is sensitive to manipulations, and can be easily activated (Williams and Eberhardt 2008). Its reduction positively influences creativity (Tadmor, Chao, Hong, and Polzer 2013). When interrupted and replaced with fluid processing styles of social categories, it lead to the adoption of more broad, inclusive and flexible representations of the social entities (Slepian, Weisbuch, Pauker, Bastian, and Ambady 2013). The latter

study, drawing on the embodied cognition paradigm, shows that simply executing fluid movements leads to less rigid processing styles, in contrast to essentialist thinking. This is applied mostly to large social categories (members of social, racial or ethnic groups) and, as Slepian and colleagues attempted to show (2013), there are few reasons to believe that these influences on processing styles also affect specific contents describing different types of personality features (person-level essentialism). Other factors, such as the socio-economic context, determine the types of theories people endorse. For instance, data from Chen, Chiu and Chan (2009) indicate that in contexts with reduced workforce mobility, people adopt more rigid views of the world.

Judging the intrapersonal and interpersonal variability of moral traits is crucial, since this evaluation serves diverse functions. Among others, the most important are the task of inferring fairness/justice and ensuring reciprocity in interactions (Baumard, Andre, and Sperber 2013; Lerner and Miller, 1978), the need for control (Kay, Gaucher, McGregor, and Nash 2010), and the need to maintain social ties (Haidt and Graham 2009). Theories regarding regularities of moral conduct are important because of their motivational and epistemic functions. These theories influence moral evaluation, as it is reflected by a research result that indicates the saliency of morally relevant concepts over moral neutral stimuli (Gantman and Van Bavel 2014); also, moral construal determines evaluation that is faster, tends to be more extreme and to lead to universal prescriptions than do non-moral evaluations of the same actions, for instance, the pragmatic and hedonic consequences (Van Bavel, Packer, Haas, and Cunningham 2012). Moral decision making can be approached, as some authors indicate, with different mindsets (Tetlock 2002). Mindsets are domain specific, but, as already stated, they are usually organized in a coherent fashion and interact with several contextual and intrapersonal variables.

Implicit Theories of Morality and Ascription of Moral Responsibility and Blameworthiness

Folk psychology is described as functioning analogously to scientific thinking, especially when it comes to describing, explaining and predicting mental states and actions (Gopnik and Wellman 1992). This perspective is best described by metaphors such as intuitive statistician or intuitive scientist, a popular representation at the dawn of the cognitivist era (Heider 1958; Kelley 1967). However, evidence regarding the vulnerability of human attribution and other person perception processes to numerous cognitive biases leads some researchers to the conclusion that this view is not just exaggerated, but inaccurate (Knobe 2010). The experimental philosopher argues that even if the set of epistemic competencies of the lay psychologist would be considered basically complete, their efficiency is affected by several factors, which disqualifies this knowledge compared with scientific cognition. Unlike academic thought, lay epistemics is characterized by many conceptual overlaps and

entanglements that contrasts the before-mentioned view of the intuitive scientist. For instance, several classes of intuitions are considered to be skewed by moral judgments (Knobe 2010). In contrast to the “theory theory” perspective, which describes persons as lay scientists, who rely on a thorough analysis of facts prior to reaching evaluations and decisions, Knobe suggests that all judgments are influenced by an instant, intuitive, moral evaluation of the action. In other words, people’s attempts at understanding the world are infused by this pervasive tendency toward moral judgment, which is done automatically and unconsciously:

Even the processes that look most “scientific” actually take moral considerations into account. It seems that we are moralizing creatures through and through (Knobe 2010, 328).

This view echoes Haidt’s (2012) position that when it comes to moral reasoning, people spend most of their time in rationalizing their intuitions or finding post-hoc justification of their decisions. This effort shows internal organization and most of the views people hold about the nature of the social world, personality and their own configuration of traits or their self-concepts are in general coherently structured (Chiu, Dweck, Tong, and Fu 1997).

When commonsense views of character are discussed in the situationist/dispositionist debate, a central issue regards the dynamics of moral behavior, or the degree to which people form, hold and endorse characterological representations both of themselves and of others. The discussion of these theories seems to affect people’s evaluations in different ways, as the research seems to reveal.

Responses to transgressions, namely culpability judgments are differently influenced by the dominant implicit theory. Entity theories describe moral behavior as reliant on stable, unchangeable traits or properties of individuals (Dweck, Chiu, and Hong 1995). In contrast, incremental theorists hold a more process-oriented view, seeing moral conduct as more malleable, and more dynamic. They claim that they achieve more accurate verdicts because they consider data and formulate more “objective” verdicts, by collecting data from multiple sources and about more diverse aspects of the case (Gervy, Chiu, Hong, and Dweck 1999). This study shows that when making attributions, entity theorists rely on internal concepts and draw conclusions from a limited amount of information (singular or very few instances); incremental theorists, on the other hand, reach their decisions by relating behavior to the situation and attend information regarding goals, reasons, transitory states, and circumstances. Also, entity theorists evaluate transgressions more harshly and propose more punitive measures. Thus, Gervy and his collaborators (1999) indicate that when evaluating responsibility, entity theorists rely to a greater degree on dispositional information about the reputation of the target. Some authors provided evidence that not only cognitive (priority over data, types of information processing, inferences, ways of reaching decisions and formulating

predictions), but also affective reactions to transgressions differentiates entity and incremental theorists, with the first being more prone to experience disposition-related emotions. For instance, Miller, Burgoon and Hall (2007) reveal that entity theorists report righteous anger and moral indignation (and other attribute related emotions), in contrast to incrementalists, who display anger and disgust, which are related to acts and specific events. Gervy, Chiu, Hong, and Dweck (1999) also underlined that entity theorists are pulled towards deciding and imposing punishment and retribution. They are motivated by vengeance more than incremental theorists, who are mainly motivated by restorative philosophies, corrective measures, education and rehabilitation. More, Chiu, Dweck, Tong, and Fu (1997) show that entity theorists are primarily motivated by conservative concerns, focused on the preservation of the status quo, while incrementalists are driven by change and social progress. These orientations, in turn, are expected to affect the priority given to specific types of information and influence decision processes related to assignment of culpability and punishment. For instance, Wurthmann (2013) describes his results on the association between implicit theories and duties versus rights-based morality, indicating that entity theorists prioritize duties and obligations as a basic frame of understanding morality, while incremental theorists endorse a rights-based perspective and consider upholding human rights as the central axis for moral evaluation. This conclusion has been reached after observing that entity theorists show a high moral awareness of violations of duties, while incremental theorists are more sensitive to violations of rights.

Franc and Tong (2014) showed that implicit theories moderated the adoption of a more transactional view of moral conduct. More specifically, entity theorists were more predisposed to praise their own future behaviors after recalling morally good or bad personal behavior, in contrast to incremental theorists, who did not display this pattern of over-glorification; in comparison, the tendency to believe in the malleability of moral character was associated to a lower tendency to commend morally praiseworthy prospective actions. Both studies indicate an underlying feature of entity theorists: the predisposition towards self-presentational concerns and strategies.

The Problem with Dweck's Model. Thinking About Traits in Positive and Negative Terms

In this section we will argue the importance of re-examining the construct of implicit theories, applied to the domain of moral conduct. Despite the accumulated evidence on the predictive power of implicit theories, other studies find no effects of implicit theories about human character evaluations, decisions and moral behavior of participants (e.g. Jessen 2014). The natural dichotomy of moral traits is an aspect that has been ignored in this research strand. People operate with concrete, either positive or negative traits, not with abstract, inert constructs (as psychological traits are treated in psychology). They employ

labels that function as moral mandates for praise or correction. Unlike other phenomena studied with this methodology (intelligence, giftedness, wisdom or willpower), the morality dimension, seen as a personal set of features, holds a peculiar kind of characteristic, the fact that lay thinking automatically attaches a valence to its descriptive terms.

The distinction between positive and negative traits or character dimensions has direct effect on the way entity and incremental theories are defined and measured. Since moral traits are conceived both on a negative (vices, moral flaws, negative traits of character) and on a positive (virtues, strengths of character) axis, the processes behind their understanding and inference could be different. These aspects can co-occur and interact, sometimes in a puzzling way for the classic personality psychology that emphasize stability and internal coherence. They need clear distinctions and workable, sensitive models. Proverbs and sayings regarding either the static or the dynamic nature of character tend to cover into both negative and positive aspects, and this double orientation tends to be influenced by cultural and religious beliefs. Entity beliefs that are expressed in folk constructions, such as: "Character is easier kept than recovered," "The wolf changes only his coat – not his character," "A leopard cannot change its spots," "Like a fence, character cannot be strengthened by whitewash," "Once a thief, always a thief" or "Give a dog a bad name and hang him." Incremental statements can be found in proverbs, covering both positive and negative aspects, including various views, such as the possibility to change for good ("It's never too late to turn over a new leaf"), and also for the bad (as in the biblical passage: "Your character can be corrupted by bad company" (1 Corinthians 15:33), and the possibility to build virtues and the self-sustaining nature of virtues ("Blood is inherited; character is earned," "Jade requires chiseling before becoming a gem," "Virtue is its own reward"). We are, at this point, interested in how the geography of the character and the mechanics of moral behavior take form in people's representations, and less concerned with the accuracy of these statements.

In this section, we will analyze some conceptual ambiguities and methodological flaws we found in Dweck and colleagues' model of implicit theories of morality. A first task in employing this model would be to separate the contents of the two alternative types of theories, in order to avoid confusion and obtain empirically testable predictions. Although the results of Dweck's team seem to be promising, at least in respect to the degree to which these theories tend to bias evaluations and generate more emotionally infused reactions, they have been less appealing to researchers in the field of morality. This could be generated by their questionable content and construct validity, mostly stemming from ambiguously defined conceptual dimensions.

We acknowledge that this model is not intended as an all-encompassing model of traits. When applied to the moral domain, which is, as we have discussed until now, dominated by some seemingly unresolvable controversies,

Dweck's conceptualization becomes oversimplifying. When Chiu, Dweck, Tong, and Fu measure the degree to which people believe that character can be shaped, starting from the premise that individuals conceive characters either as fixed, or as fluid, malleable, they allow people to apply various personal definitions of character.

Keeping in mind Meindl and Graham's (2013) concept of lay moral disagreement and Kohlberg's (1972) view about the variability of understanding morality traits, we believe that, in absence of other clarifications, these items cannot provide workable results. The reason why moral valence is neglected is not clear. The items used to measure implicit theories do not necessarily express the view of the educability and malleability of positive moral traits, as incremental views intend, but they may very well depict a belief in what Zimbardo referred to as the subtle and gradual dehumanization of people or the incremental nature of evil. As Aquinas put it, "the evil of mutable spirits [i.e., human beings] arises from the evil choice itself, and that evil diminishes and corrupts the goodness of nature". Chiu, Dweck, Tong, and Fu (1997) obtained a composite measure subtracted from the scores on the implicit theories of morality and of the world. Based on this, they decided to categorize the study participants, a decision that contaminates the validity of this study. This decision has been assumed and perpetuated by other researchers using this methodology.

Although Chiu, Dweck, Tong, and Fu define incremental theorists as those who "believe that these factors can be shaped, cultivated, or improved" (Chiu, Dweck, Tong, and Fu 1997, 924), the way they assign subjects to this category is problematic. They chose to use only items that reflect an entity view of people's morality ("A person's moral character is very basic and cannot be changed very much") and decided that low scorers automatically endorse the alternative view. Despite this bipolar structure, they claimed that the construct is "unidimensional and defined by a unitary idea" (Chiu, Dweck, Tong, and Fu 1997, 925), a feature that they considered it justifies their decision to exclude items depicting incremental theories in the survey. This decision was based on the claim that, in the pilot studies, incremental items proved to be compelling for respondents, offering a skewed distribution. However, these items are imprecise ("The basic moral characteristic of a person can be changed significantly, no matter who this person is;" "Even the most basic moral qualities of a person can be changed") and require adjusting their contents with more specific terms, referring to various components of human character.

Based on these considerations, we propose a revision of the content validity of the concept of "implicit theories of morality." We intend to refine its dimensionality by introducing the items that were dispelled based on psychometric reasons but which invalidate all the claims regarding the assignment of subjects into entity and incrementalist theories. As already stated, the concepts of character and traits are disputed and fragmented. Even when used in their traditional sense in psychology, namely as diagnostic tools, they

have some common feature: they are multifaceted and charged with moral valence.

This dichotomy comprised in the evaluative representations of character attributes is not adequately caught by the description and measures of implicit theories (as proposed by Chiu Dweck, Tong, and Fu 1997), whose generality avoids the evaluative aspect, and offers thus incomplete and ambiguous descriptions of the central traits. The items are the following three: *A person's moral character is something very basic about them, and it can't be changed much; Whether a person is responsible or sincere or not is deeply ingrained in their personality. It cannot be changed very much; and There is not much that can be done to change a person's moral traits.* Usually, participants are asked to indicate their agreement on a 6-point Likert scale (ranging from 1 - very strongly agree, to 6 - very strongly disagree). The questions force the respondents to fit all these aspects in a presumably all-encompassing notion of moral attributes, and this choice implicitly suggests that trait-related properties are similar for all the involved dimensions, either positive or negative. As Zimbardo pointed out, incrementalism can be applied to the development of "evil" traits, negative habits or vices and the first and third items can be interpreted in either the positive or the negative direction of the moral compass. Also, items the first two items cover more aspects, specific to the facets of the characterological view, including innateness of good and bad moral features, educability or, on the contrary, dissolubility of virtues.

In other words, intra-individual variability and contextual sensitivity of traits (cross-situational flexibility) are forcibly treated equally, as a consequence of the wording of items. Another problem of this instrument comes from the authors' decision to eliminate the items tapping the incremental views. Chiu, Dweck, Tong, and Fu (1997) and Gervery, Chiu, Hong, and Dweck (1999) argue that the items are socially desirable and maintain that disagreement with an entity theory automatically can be interpreted as an endorsement of the incremental theory. The authors do not provide convincing reasons for discharging incremental view related items and we suspect that their choice was based on the need to offer the same type of measure as those already applied to other descriptors (e.g. intelligence, creativity, wisdom). Thus, the concept and its corresponding cognitive measure should have a more veridical reference and precise content, without forcing into the same assessment diverging or even contradictory evaluative information.

The separation between "good" and "bad" positive traits and between virtues and vices is necessary and cannot be ignored. The controversial semantics of the items just discussed calls for a refinement of the methodology used for implicit theories of morality. We propose the segregation of information representations on two categories or trait domains: strengths vs. weaknesses or virtues vs. vices. When subjects read the item *"A person's moral character is something very basic about them, and it can't be changed much"*, some might

think of virtues, while others could have vices in mind. The concept of character is multifaceted and includes sets of traits, motives and competencies, such as self-control, willpower, adherence to moral desires, values and principles, consistency, and integrity. Some moral judgments, behaviors or attitudes rely on traits governed by spontaneous processes, while other are more deliberative in nature, and thus are characterized by variable amounts of fixedness or educability.

Furthermore, if all these types can be conceptualized as modifiable or buildable, the same incremental principle can be applied to vices but it doesn't have to vary in a similar way. In addition to that, once acquired, the virtue or vice can be viewed as stable or degradable and we suspect again that this is not necessarily due to the fact that they show the same properties or follow similar principles. From this point of view, the incremental and entity principles are contrary but not necessarily contradictory. They can be logically, but not necessarily psychologically incompatible and this dissociation has to do more with the complexity of the moral character than with the limits of human rationality. This calls for building a more refined set of measures of implicit theories of morality that reflect the complexity of the domain. People can endorse items from both categories (logically incompatible), holding in mind different dimensions of character. For instance, they can consider self-control or empathy as a fixed entities, while considering expertise in moral judgment, or values as malleable aspects.

The one-dimensional model of implicit theories is too restrictive for understanding moral concepts and traits. Taking valence into account is crucial. We do not intend to prove that some traits are normatively blameworthy and other praiseworthy, but that they are charged with good/bad, positive/negative representational content.

We agree with Schwarz and Sharpe's (2010) argument that traits are not inherently good or bad, with the exception of wisdom, which allows the individual to decide how to act in particular circumstances. As in the case of wisdom, other authors describe "moral expertise" as the set of abilities that allow the accurate apprehension of moral situations, the capacity to reach the adequate solution and also to carry the appropriate act, that ultimately is defining for one's self-concept and becomes embedded in the person's moral identity (Husley and Hamson 2014; Narvaez and Lapsley 2005). Husley and Hamson (2014) develop a model that starts from the Thomistic concept of habitus that represents the idea that acting on moral beliefs consolidates these beliefs and that further increases the likelihood of acting according to them. These habits are dispositions towards patterns of feeling and acting, according to a certain moral identity, or to strive towards these moral goals or ideals. Taking into account the automaticity (often interpreted by moral psychologists as intuitive nature) of moral evaluation, the authors claim that this process is similar to the formation of habits.

Lay moral disagreement indicates that the same behavior can be interpreted as indicative of a positive or negative trait, as a function of the context, of the perceiver's beliefs and of induced mindsets. Thus, another step is to supplement the instrument with a scale that could measure whether people, as intuitive ethicists, are, in the terms of a classical debate in analytic moral philosophy, particularists or generalists. Are they recognizing universal, defensible and fundamental moral principles, in an Aristotelian and Kantian sense, or on the contrary, disregard them, as particularists do (Dancy 1983; McDowell 1979; McNaughton 1988). According to the latter view, moral behavior is judged only in relation to specific situations. Moral evaluation cannot be based on standards and procedures, a position that entails that moral pedagogy cannot rely on codes of ethics, but only on the practical experience of the individual.

As particularists regard principles as „crutches that a morally sensitive person would not require, and indeed the use of such crutches might even lead us into moral error” (Dancy 2004), some psychologists draw attention to the risk of treating moral traits in an isolated, decontextualized fashion. Positive or negative traits can have both positive and negative outcomes in different contexts, either cultural or interpersonal (McNulty and Fincham 2012). McNulty and Fincham (2012) underline that the degree to which a behavior or display of a trait is conceived as right or wrong, good or bad, syntonic or dystonic, depends on its situational appropriateness. In an attempt to bring attention to the dangers of positive psychology's enthusiasm to build positive traits indiscriminately, as Lyubomirsky (2012) says, they argue “that forgiveness and optimism can have both beneficial and adverse consequences, depending on the context.” (Lyubomirsky 2012, 574). However, we believe that people attach two types of evaluative valence to trait labels or lay moral concepts: on the one hand, an intrinsic quality, which is trans-situational and can be an inclination to either good or bad, right or wrong behavior) and a more flexible component, that takes into account the situational adequacy of trait-relevant or trait-diagnostic behaviors, that can have different meanings, from a situation to another. An argument for this property of moral traits is Lapsley and Lasky's (2001) study on moral prototypes. They identified a list of character-central (e.g. kind, sincere, and loyal) and character-peripheral traits (lucky, energetic, and independent) of the moral prototypes, generated by their participants. The peripheral traits could be more sensitive to the situational prioritization of moral concerns, thus gaining or losing diagnosticity and moral “currency.”

Other studies provide evidence that theories of morality can function incrementally and be subjected to ontogenetic progress. In an early study, Heyman and Dweck (1998) point to the fact that 7-8 years old children have the tendency to see character as more stable than other aspects of personality, such as ability-related dimensions. However, incrementality seems to be differently related to good or positive and bad traits. Children have a tendency to display

developmental optimism, as Lockhart, Chang and Story (2002) showed. From early development stages, children have a propensity to view positive features, both physical and psychological (including moral traits) as invariant, while the negative ones are considered susceptible to gradual fading or even dissolution. In other words, developmental optimism refers to a tendency to believe that physical and moral features are subjected to influences that restores them toward more positive states. Newman, Bloom and Knobe (2014) show that people attach a fundamentally positive quality to all essential features, in the sense that vicious or bad elements are seen as transient, ephemeral, since virtues are considered the „true nature” or essence of things, or true selves. Subjects in the series of studies conducted by Newman and colleagues (2014) were asked to rate situations in which people underwent behavioral changes from good to bad and from bad to good. For each type of change, they were asked to rate whether this modification was a reflection of their true self, and the degree to which it reflected the “true to the deepest, most essential aspects of (the agent’s) being” (Newman, Bloom, and Knobe 2014, 204). The data suggest that morally good or praiseworthy behavioral modifications were generally considered as inherent to the true nature of people: “deep inside every individual there is a ‘true self’ motivating him or her to behave in ways that are virtuous” (Newman, Bloom, and Knobe 2014, 212).

Another argument for the segregation of good and bad character traits comes from studies that have shown that people do not process positive and negative character traits similarly. For instance, Fosatti and colleagues (2004) indicate that processing negative traits prompts more brain activation than processing positive descriptors. They suggest that the recollection of negative stimuli is more important than that of positive stimuli as a survival mechanism.

Another possibility that could be explored is whether not only the valence, but the specific content of moral traits, are sources in variations of incrementality judgments. For instance, it is possible that different traits, such as the ones described by moral foundations model (MFT), with the following six aspects: care/harm for others, fairness/cheating, justice, liberty/oppression, loyalty/betrayal, authority/subversion sanctity/degradation (Haidt, Graham, and Joseph 2009), could be conceived by lay perceivers as following different internal rules of functioning, as the authors of the model advocate that they rely on distinct psychological systems. The authors of the MFT model have consistently shown that liberals, for instance, do not recognize all these dimensions as relevant to the moral domain (for more details, see Haidt 2012). Given the variability of what counts as moral or not, due to cultural, interpersonal and sometimes intrapersonal (DeScioli, Massenkoff, Shaw, Petersen and Kurzban, 2014 show that self-interest can dictate endorsement of moral principles) differences, maybe more refined measurement of the implicit theories should be designed and employed in further studies.

If our proposal and attached arguments are correct, then positive and negative traits should be judged differently in terms of modifiability, stability or educability. It is plausible to find low or lacking correlations between the ascriptions of the principle of incremental change in moral character, when separating positive or negative features.

The authors of the model argue that the theories are implicit, tacit, but can be elicited by direct questions (Dweck 1999). However, this recommendation runs a risk repeatedly revealed in experimental philosophy, a field that aspires to compensate the habit of armchair philosophers “to place their own intuitions into the mouths of the folk in a way that supports their own position – neglecting to verify whether their intuitions agree with what the majority of non-philosophers actually think” (Nahmias, Morris, Nadelhoffer, and Turner 2005, 562). At the same time, there is also a parallel effect, given by the poor collaboration between psychologists and ethicists that leads to a poor grasping of concepts. There is also the practice of adapting models from one area to another, oversimplifying a complex domain such as moral appraisal. So, researchers need to take a closer look to the content and dimensionality of this supposedly “unidimensional and defined by a unitary idea” construct of implicit theories applied to the domain of morality (Chiu, Dweck, Tong, and Fu 1997, 925). We think we need a new model of scales that takes into account the differences between the incremental nature of positive and negative traits, in both positive (the dissolution of vices and the building or strengthening the virtues) and negative direction (evolution of undesirable habits and behaviors and the degradation of virtues). Revising the theoretical validity of the concept is important. This could be made first by extracting emic definitions of good/bad traits, virtues or vices from lay subjects and then the specification of the model. We expect to discover various ways in which the concept unfolds on lay discourse, for instance: innateness or genetic determinism of good traits, character strengths and habits, innateness of the proneness towards immoral behavior, malleability or educability of virtues, instability or degradability of virtues (reversibility), controllability of vices or character flaws (dispersibility), and the incremental nature of vices. These new measures will be related to specific criteria (evaluation of the predictions people make based on present evaluation, that is, how lay people estimate the patterns of change and stability, on both positive and negative evaluative dimensions).

The conceptual and methodological solution to these issues could lead to a more accurate understanding of how people think of the organization and dynamics of moral features in individuals and groups. We also hope to bring light upon some of the inconsistent findings regarding the determinants and consequences of implicit theories.

Implications of the Study of Implicit Theories of Morality

Our question is related to how people think about traits that they so readily employ in everyday interactions and to what extent moral traits are permeable to external factors. At this point, we are neither interested in debunking the components and constraints that affect naïve moral psychology, nor in evaluating their accuracy. We only approach folk conceptions of morality with the attitude reflected in the following description of the relationship between scientific knowledge and folk conceptions of the world:

whatever may be learned about folk science will have no relevance to the pursuit of naturalistic inquiry into the topics that folk science addresses in its own way (Chomsky 1994, 14).

This direction of research will ultimately inform the question of *how we make moral judgments*, and also the answer to *how we could improve our judgments*. Uncovering the processes of “judging” saints, heroes, sinners or villains and the specific enablers of the drifts from the sinning saint to the saintly sinner state and the reverse has several implications for social cognition and also for moral education.

First we need to explore the literature on moral judgments on how thinking about how morality and moral traits operate within the person affect the shape of moral cognitions and even behavior. Secondly, we need to test whether moral reasoning is malleable and whether folk or naive psychologies are responsive and permeable to new information. We could do this by changing the valence or the intensity of assessments of responsibility, biasing them towards more positivity, or, on the contrary, towards more negativity. Recognizing and correcting biases in moral folk psychology is crucial and would have an impact on many practical domains where moral judgments are inherent, such as justice or governance. This is similar to Harman’s recommendations regarding physics:

Ordinary untrained physical intuitions are often in error. (...) This means, among other things, that bombardiers need to be trained to go against their own physical intuitions. (Harman 1999, 315).

Debiasing lay perceivers or reducing the distortions brought by implicit person theories could ultimately become a goal of an understudied area in moral psychology.

Some progresses have been made in this direction. Yeager, Trzesniewski and Dweck (2013) implemented an intervention which reduces the consequences of bullying and victimization. They attempted to shape adolescents’ perceptions regarding the person’s potential to change. The majority of adolescents endorse entity theories and have a tendency to categorize people into typologies such as victims, aggressors, winners or losers (Yeager and Miu 2011; Yeager, Trzesniewski, Tirri, Nikelainen, and Dweck 2011). Entity theorists were oriented toward vengeance (in other words they

prepare a hostile, aggressive answer for the transgressor, are more inclined to make hostile attributions), while incremental theorists are more inclined to choose prosocial, educational solutions (Yeager, Miu, Powers and Dweck 2013). Interventions described by these authors reduced hostile reactions and promoted more prosocial responses to norm violations.

As we have already noted, empirical approaches in moral psychology cannot be discussed independently of some central philosophical conceptual debates. Thus, analyzing the naive conceptualizations of character would be difficult without discussing the ontological status of traits or the components of moral character (how we ought to think about character). Although we are interested in the form and contexts of implicit theories of morality, not necessarily their accuracy, sophistication and elaboration, we are optimistic that philosophically informed empirical approaches of folk morality can benefit both areas of research. For instance, the debate between situationists and dispositionists may shed light on whether the instruments employed for measuring implicit theories of personality are accurate. One of the most often critique of situationism is that it mistakes the philosophical concept of virtue to a different view of character traits. According to this view, traits are rigid, associated with habitual or stereotypical patterns of behavior, ignoring the reflective, incremental quality of virtue. Other problems arise from the poor definitions that philosophers attach to their concepts: for instance, are these traits features of normal, psychologically healthy persons or are they ideals, prototypes of heroes, saints or other virtuous people.

Research should also explore whether the evaluative valences of traits are stable and whether this proposed dichotomy (on the axiological dimension – good/ bad or on the normative one – appropriate/ inappropriate or right/wrong) stand the empirical test. Previous research focused on convergent validation of implicit theories, and on their predictive validity. There are scarce attempts at finding suitable criteria for demonstrating that responses on the items designed to measure implicit theories of morality relate to the actual tendency toward making dispositional inferences. These new measures will be related to criteria such as evaluation of the predictions people make based on present evaluation, how lay people estimate the patterns of change and stability, on both positive and negative evaluative dimensions. Also, the issue of temporal stability is inadequately explored. Several state manipulations or primes indicate high sensitivity of the aspects but do not say much about the chronic resilience of the implicit theories. We believe that all these issues merits further theoretical debate and experimentation from philosophers and psychologists interested in both person perception, justice reasoning and moral appraisal.

References

Alfano, Mark. 2013. *Character as Moral Fiction*. Cambridge: Cambridge University Press.

Ana Maria Hojbotă

- Allport, Gordon Willard. 1937. *Personality: A Psychological Interpretation*. Oxford: Holt.
- Anscombe, Gertrude E. M. 1958. "Modern Moral Philosophy." *Philosophy* 33: 1-19.
- Aristotle. *The Complete Works of Aristotle: The Revised Oxford Translation*, ed. Jonathan Barnes. Princeton: Princeton University Press, 1984.
- Bastian, Brock and Nick Haslam. 2006. "Psychological Essentialism and Stereotype Endorsement." *Journal of Experimental Social Psychology* 42: 228-235.
- Baumard, Nicolas, Jean-Baptiste André, and Dan Sperber. 2013. "A Mutualistic Approach to Morality." *Behavioral and Brain Sciences* 36 (1): 59-122.
- Callan, Mitch J., Robbie M. Sutton, Annelie J. Harvey, and Rael J. Dawtry. 2014. "Immanent Justice Reasoning: Theory, Research, and Current Directions." In *Advances in Experimental Social Psychology*, vol. 49, eds. J.M. Olson and M.P. Zanna, 105-161. London: Academic Press.
- Chen, Jing, Chi-yue Chiu, and Fiona S. Chan. 2009. "The Cultural Effects of Job Mobility and the Belief in a Fixed World: Evidence from Performance Forecast." *Journal of Personality and Social Psychology* 97: 851-865.
- Chiu, Chi-yue, Carol S. Dweck, Jennifer Y. Y. Tong, and Jeanne H. Y. Fu. 1997. "Implicit Theories and Conceptions of Morality." *Journal of Personality and Social Psychology* 73: 923-940.
- Chomsky, Noam. 1995. "Language and Nature." *Mind* 104 (431): 1-61.
- Church, A. Timothy, Fernando A. Ortiz, Marcia S. Katigbak, Tatyana V. Avdeyeva, Alice M. Emerson, Jose de Jesus F. Vargas, and Joselina R. Ibanez. 2003. "Measuring Individual and Cultural Differences in Implicit Trait Theories." *Journal of Personality and Social Psychology* 85: 332-347.
- Dancy, Jonathan. 1983. "Ethical Particularism and Morally Relevant Properties." *Mind* 92 (368): 530-547.
- Dancy, Jonathan. 2004. *Ethics Without Principles*. Oxford: Oxford University Press.
- DeScioli, Peter, Maxim Massenkoff, Alex Shaw, Michael Petersen, and Robert Kurzban. 2014. "Equity or Equality? Moral Judgments Follow the Money." *Proceedings of the Royal Society B* 281: 1-6.
- DeSteno, David A. and Piercarlo Valdesolo. 2011. *Out of Character: Surprising Truths about the Liar, Cheat, Sinner (and Saint) lurking in all of us*. New York: Crown.
- Doris, John. 2002. *Lack of Character: Personality and Moral Behavior*. New York: Cambridge.
- Dweck, Carol S. 1991. "Self-theories and Goals: Their Role in Motivation, Personality, and Development." In *Nebraska Symposium on Motivation*, ed. R. Dienstbier. Lincoln, Nebraska: University of Nebraska Press.
- Dweck, Carol S. 1999. *Self Theories: Their Role in Motivation, Personality, and Development*. Hove: Psychology Press.

- Implicit Theories of Morality, Personality, and Contextual Factors in Moral Appraisal
- Dweck, Carol S., Chi-Yue Chiu, and Ying-yi Hong. 1995. "Implicit Theories and their Role in Judgments and Reactions: A World from Two Perspectives." *Psychological Inquiry* 6: 267–285.
- Easterbrook, Gregg. 2001. "March 5. I'm OK, you're OK." *The New Republic*: 20–23.
- Emmons, Robert A. 2009. "Greatest of the Virtues? Gratitude and the Grateful Personality." In *Personality, Identity, and Character: Explorations in Moral Psychology*, ed. D. Narvaez and D. K. Lapsley, 256–270. New York: Cambridge University Press.
- Flanagan, Owen. 1991. *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, MA: Harvard University Press.
- Fossati, Phillippe, Stephanie J. Hevenor, Martin Lepage, Simon J. Graham, Cheryl Grady, Michelle L. Keightley, Fergus Craick, and Helen Mayberg. 2004. "Distributed Self in Episodic Memory: Neural Correlates of Successful Retrieval of Self-encoded Positive and Negative Personality Traits." *Neuroimage* 22: 1596–1604.
- Franc, Michael and Yuk-Yue Tong. 2014. "Implicit Theories of Morality and Moral Licensing." Paper presented at the Association for Psychological Science 2014 convention in San Francisco.
- Furnham, Adrian. 2003. "Belief in a Just World: Research Progress over the Past Decade." *Personality and Individual Differences* 34: 795–817.
- Gantman, Ana P. and Jay J. Van Bavel. 2014. "The Moral Pop-out Effect: Enhanced Perceptual Awareness of Morally Relevant Stimuli." *Cognition* 132: 22–29.
- Geertz, Clifford. 1984. "From the Native's Point of View: On the Nature of Anthropological Understanding." In *Culture Theory*, ed. R. Shweder and R. LeVine, 123–136. Cambridge, UK: Cambridge University Press.
- Gervey, Benjamin, Chi-yue Chiu, Ying-yi Hong, and Carol. S. Dweck. 1999. "Differential Use of Person Information in Decision-making about Guilt vs. Innocence: The Role of Implicit Theories." *Personality and Social Psychology Bulletin* 25: 17–27.
- Gill, Michael. 2009. "Indeterminacy and Variability in Meta-Ethics." *Philosophical Studies* 145(2): 215–234.
- Gopnik, Alison and Henry M. Wellman. 1992. "Why the Child's Theory of Mind Really is a Theory." *Mind and Language* 7 (1-2): 145–171.
- Graham, Jesse, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Petter H. Ditto. 2013. "Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism." *Advances in Experimental Social Psychology* 47: 55–130.
- Greene, Joshua. 2013. *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. New York: Penguin Press.
- Guglielmo, Steve, Andrew E. Monroe, and Bertram F. Malle. 2009. "At the Heart of Morality Lies Folk Psychology." *Inquiry* 52: 449–66.

- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York, NY: Pantheon.
- Haidt, Jonathan and Jesse Graham. 2009. "Planet of the Durkheimians, where Community, Authority, and Sacredness are Foundations of Morality." In *Social and Psychological Bases of Ideology and System Justification*, ed. J. Jost, A. C. Kay, and H. Thorisdottir, 371–401. New York: Oxford University Press.
- Haidt, Jonathan, Jesse Graham, and Craig Joseph. 2009. "Above and Below Left-Right: Ideological Narratives and Moral Foundations." *Psychological Inquiry* 20: 110-119.
- Hamlin, J. Kiley, Karen Wynn, and Paul Bloom. 2007. "Social Evaluation by Preverbal Infants." *Nature* 450: 557–60.
- Harman, Gilbert. 1999. "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error." *Proceedings of the Aristotelian Society* 99: 315- 331.
- Haslam, Nick, Brock Bastian, Paul Bain, and Yoshi Kashima. 2006. "Psychological Essentialism, Implicit Theories, and Intergroup Relations." *Group Processes & Intergroup Relations* 9: 63-76.
- Heider, Fritz. 1958. *The Psychology of Interpersonal Relations*. New York: Wiley.
- Heider, Fritz, and Marianne Simmel. 1944. "An Experimental Study of Apparent Behavior." *The American Journal of Psychology* 57(2): 243–259.
- Heyman, Gail D., and Carol S. Dweck. 1998. "Children's Thinking about Traits: Implications for Judgments of the Self and Others." *Child Development* 64: 391-403.
- Hulsey, Timothy L. and Peter J. Hampson. 2014. "Moral Expertise." *New Ideas in Psychology* 34: 1-11.
- Jessen, Annika. 2014. "The Effect of Perceived Intentions to Harm and Implicit Theories of Personality on Moral Judgments of Favoritism and Discrimination". CMC Senior Theses. Paper 829. At http://scholarship.claremont.edu/cmcs_theses/829.
- Jones, Edward E. 1990. *Interpersonal Perception*. New York: W. H. Freeman and Company.
- Kamtekar, Rachana. 2004. "Situationism and Virtue Ethics on the Content of Moral Character." *Ethics* CXIV: 458–491.
- Kay, Aaron C., Danielle Gaucher, Ian McGregor, and Kyle Nash. 2010. "Religious Belief as Compensatory Control." *Personality and Social Psychology Review* 14: 37–48.
- Keller, Johannes. 2005. "In Genes We Trust: The Biological Component of Psychological Essentialism and Its Relationship to Mechanisms of Motivated Social Cognition." *Journal of Personality and Social Psychology* 88: 686-702.

- Implicit Theories of Morality, Personality, and Contextual Factors in Moral Appraisal
- Kelley, Harold H. 1967. "Attribution Theory in Social Psychology". In *Nebraska Symposium on Motivation*, ed. D. Levine, 192-238. Lincoln: University of Nebraska Press.
- Knobe, Joshua. 2010. "Person as Scientist, Person as Moralist." *Behavioral and Brain Sciences* 33: 315-329.
- Kohlberg, Lawrence and Rochelle Mayer. 1972. "Development as the Aim of Education." *Harvard Educational Review* 42: 449-496.
- Korsgaard, Christine M. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Kupperman, Joel J. 2009. "Virtue in Virtue Ethics." *Journal of Ethics* 13: 243-255.
- Lapsley, Daniel K., and Benjamin Lasky. 2001. "Prototypic Moral Character." *Identity: An International Journal of Theory and Research* 1: 345-363.
- Lerner, Melvin J., and Dale T. Miller. 1978. "Just World Research and the Attribution Process: Looking back and ahead." *Psychological Bulletin* 85: 1030-1051.
- Levy, Sheri R., Steven J. Stroessner, and Carol S. Dweck. 1998. "Stereotype Formation and Endorsement: The Role of Implicit Theories." *Journal of Personality and Social Psychology* 74: 1421-1436.
- Leyens, Jaques-Phillippe, Paola M. Paladino, Ramon T. Rodriguez, Jeroen Vaes, Stephanie Demoulin, Armando P. Rodriguez, and Ruth Gaunt. 2000. "The Emotional Side of Prejudice: The Attribution of Secondary Emotions to Ingroups and Outgroups." *Personality and Social Psychology Review*, 4: 186-197.
- Lockhart, Kristi L., Bernard Chang, and Tyler Story. 2002. "Young Children's Beliefs about the Stability of Traits: Protective Optimism?" *Child Development* 73: 1408-1430.
- Lyubomirsky, Sonja. 2012. "Positive Psychologists on Positive Constructs." [Comment on McNulty and Fincham's "Beyond Positive Psychology? Toward a Contextual View of Psychological Processes and Wellbeing"]. *American Psychologist* 67: 574.
- McAdams, Dan P. 2009. "The Moral Personality." In *Personality, Identity, and Character: Explorations in Moral Psychology*, ed. D. Narvaez and D. K. Lapsley, 11-29. New York: Cambridge University Press.
- McDowell, John. 1979. "Virtue and Reason." *The Monist* 62: 331-350.
- McNaughton, David. 1988. *Moral Vision*. Oxford: Blackwell.
- McNulty, James K., and Frank D. Fincham. 2012. "Beyond Positive Psychology? Toward a Contextual View of Psychological Processes and Wellbeing." *American Psychologist* 67: 101-110.
- Meindl, Peter and Jesse Graham. 2014. "Know Thy Participant: The Trouble with Nomothetic Assumptions in Moral Psychology." In *Advances in Experimental Moral Psychology*, ed. H. Sarkissian and J. C. Wright, 233-252. New York: Continuum Press.

- Miller, Claude H., Judee K. Burgoon, and John R. Hall. 2007. "The Effects of Implicit Theories of Moral Character on Affective Reactions to Moral Transgressions." *Social Cognition* 25: 819-832.
- Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer, and Jason Turner. 2005. "Surveying Freedom." *Philosophical Psychology* 18 (5): 561-584.
- Narvaez, Darcia and Daniel Lapsley. 2005. "The Psychological Foundations of Everyday Morality and Moral Expertise." In *Character Psychology and Character Education*, ed. D. Lapsley and C. Power, 140-165. Notre Dame: IN: University of Notre Dame Press.
- Newman, George E., Paul Bloom, and Joshua Knobe. 2014. "Value Judgments and the True Self." *Personality and Social Psychology Bulletin* 40 (2): 203-216.
- Nisbett, Richard E., Kaiping Peng, Incheol Choi, and Ara Norenzayan. 2001. "Culture and Systems of Thought: Holistic vs. Analytic Cognition." *Psychological Review* 108: 291-310.
- Nussbaum, Martha C. 1986. *The Fragility of Goodness. Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge: Cambridge University Press.
- Osswald, Silvia, Tobias Greitemeyer, Peter Fischer, and Dieter Frey. 2010. "Moral Prototypes and Moral Behavior: Specific Effects on Emotional Precursors of Moral Behavior and on Moral Behavior by the Activation of Moral Prototypes." *European Journal of Social Psychology* 40 (6): 1078-1094.
- Peterson, Christopher and Martin Seligman. 2004. *Character Strengths and Virtues: A Handbook and Classification*. Washington, DC: APA Press.
- Plaks, Jason. E., Steven. J. Stroessner, Carol. S. Dweck, and Jeffrey. W. Sherman. 2001. "Person Theories and Attention Allocation: Preferences for Stereotypic vs. Counter-stereotypic Information." *Journal of Personality and Social Psychology* 80: 876-893.
- Rai, Taze S. and Alan P. Fiske. 2011. "Moral Psychology is Relationship Regulation: Moral Motives for Unity, Hierarchy, Equality, and Proportionality." *Psychological Review* 118: 57-75.
- Ross, Lee and Richard E. Nisbett. 1991. *The Person and the Situation: Perspectives of Social Psychology*. Philadelphia: Temple University Press.
- Sachdeva, Sonya, Rumén Iliev, and Douglas L. Medin. 2009. "Sinning Saints and Sainly Sinners: The Paradox of Moral Self-Regulation." *Psychological Science* 20 (4): 523-528.
- Sarkissian, Hagop, John Park, David Tien, Jennifer Wright, and Joshua Knobe. 2012. "Folk Moral Relativism." *Mind and Language* 26 (4): 482-505.
- Schwartz, Barry, and Kenneth Sharpe. 2010. *Practical Wisdom: The Right Way to Do the Right Thing*. New York: Riverhead.
- Shoda, Yuichi, Scott L. Tiernan, and Walter Mischel. 2002. "Personality as a Dynamical System: Emergence of Stability and Distinctiveness from Intra- and Interpersonal Interactions." *Personality and Social Psychology Review*, 6: 316-25.

Implicit Theories of Morality, Personality, and Contextual Factors in Moral Appraisal

- Slepian, Michael L., Max Weisbuch, Kristin Pauker, Brock Bastian, and Nalini Ambady. 2014. "Fluid Movement and Fluid Social Cognition: Bodily Movement Influences Essentialist Thought." *Personality and Social Psychology Bulletin* 40: 111-120.
- Statman, Daniel. 1997. "Introduction to Virtue Ethics." In *Virtue Ethics: A Critical Reader*, ed. Daniel Statman, 1-42. Edinburgh: Edinburgh University Press.
- Tadmor, Carmit T., Melody M. Chao, Ying-yi Hong, and Jeffrey, T. Polzer. 2013. "Not Just for Stereotyping Anymore: Racial Essentialism Reduces Domain-General Creativity." *Psychological Science* 24: 99-105.
- Tetlock, Phillip E. 2002. "Social-Functionalist Metaphors for Judgment and Choice: The Intuitive Politician, Theologian, and Prosecutor." *Psychological Review* 109: 451-472.
- Van Bavel, Jay J., Dominic J. Packer, Ingrid J. Haas, and William A. Cunningham. 2012. "The Importance of Moral Construal: Moral versus Non-Moral Construal Elicits Faster, More Extreme, Universal Valuations of the Same Actions." *PLoS ONE*, 7, e48693. 10.1371/journal.pone.0048693.
- Verkuyten, Maykel. 2003. "Discourses about Ethnic Group (De-) Essentialism: Oppressive and Progressive Aspects." *British Journal of Social Psychology* 42: 371-391.
- Williams, Melissa J. and Jennifer L. Eberhardt. 2008. "Biological Conceptions of Race and the Motivation to Cross Racial Boundaries." *Journal of Personality and Social Psychology* 94: 1033-1047.
- Wurthmann, Kurt. 2013. "Perceivers' Implicit Theories, Violation Type and Moral Awareness." *Academy of Management Proceedings* 772.
- Zimbardo, Phillip. 2007. *The Lucifer Effect: Understanding How Good People Turn Evil*. New York: Random House.
- Yeager, David S. and Adriana Miu. 2011. "Implicit Theories of Personality Predict Motivation to Use Prosocial Coping Strategies after Bullying in High School." In *Personality, Stress and Coping: Implications for Education*, eds. E. Frydenberg and G. Reeve, 47-62. Charlotte: Information Age.
- Yeager, David S., Adriana Miu, Joseph Powers, and Carol S. Dweck. 2013. "Implicit Theories of Personality and Attributions of Hostile Intent: A Meta-Analysis, an Experiment, and a Longitudinal Intervention." *Child Development* 84: 1651-1667.
- Yeager, David S., Kali Trzesniewski, and Carol S. Dweck. 2013. "An Implicit Theories of Personality Intervention Reduces Adolescent Aggression in Response to Victimization and Exclusion." *Child Development* 84: 970-988.
- Yeager, David S., Kali H. Trzesniewski, Kirsi Tirri, Petri Nokelainen, and Carol S. Dweck. 2011. "Adolescents' Implicit Theories Predict Desire for Vengeance After Remembered and Hypothetical Peer Conflicts: Correlational and Experimental Evidence." *Developmental Psychology* 47: 1090-1107.