

The role of morality in social cognition

Jennifer L. Ray
New York University

Peter Mende-Siedlecki
University of Delaware

Ana P. Gantman
City University of New York

Jay J. Van Bavel
New York University

Ray, J. L., Mende-Siedlecki, P., Gantman, A. P., & Van Bavel, J. J. (in press). The role of morality in social cognition. In K. Ochsner & M. Gilead. (Eds.) *The Neural Bases of Mentalizing*. Springer Press.

Abstract

Over the past few decades, two-factor models of social cognition have emerged as the dominant framework for understanding impression formation. Despite the differences in the labels, there is wide agreement that one dimension reflects sociability potential, and the other, competence. One way in which the various two-factor models *do* clearly differ, however, is in the way the dimensions incorporate or produce evaluations of morality. Aristotle saw morality as the most important basis on which to form positive evaluations, because competence and sociability could only be virtuous, sincere, and trustworthy if expressed through a moral character. This chapter highlights research demonstrating the unique and possibly primary role of morality in social cognition. We clarify the dynamic, interactive, and conjoint effects of morality on social perception, and argue morality, competence, and sociability are three influential and interactive dimensions of social perception.

Keywords: social cognition, impression formation, morality, dynamical systems

The role of morality in social cognition

Over the past few decades, two-factor models of social cognition have emerged as the dominant framework for understanding impression formation. Despite the differences in the labeling of the two factors, there is wide agreement that the core of one dimension reflects social/relational potential (which we will call sociability), and the other, competence/capacity (which we will call competence). However, scholars dating back to Aristotle have argued that morality may be the most important basis on which to form social evaluations, because competence and sociability could only be virtuous, sincere, and trustworthy if expressed through a moral character (MacIntyre, 1984). Indeed, recent work demonstrates that morality judgments influence the evaluation of other characteristics, shaping evaluations of the two core dimensions that dominate the literature in person perception: competence and sociability. In the current chapter, we will: (1) briefly describe several two-factor models of social cognition, (2) review evidence for why morality should be treated as a unique and primary dimension, and (3) discuss the flexibility of impression formation due to goals that affect attention to, and appraisal of, the moral dimension. This review reflects a growing consensus that two-factor models fail to capture the rich nature of impression formation (Tamir, Thornton, Contreras, & Mitchell, 2016; Thornton & Mitchell, 2018) and that morality might be the most important dimension of person perception (Goodwin, 2015).

Models of Social Cognition

Although mentalizing is often described as thinking about transient mental states, there is a close, bi-directional relationship between inferring mental states and attributing stable traits. Indeed, some have argued that personality ought to be reconceptualized in terms of repeated action patterns and tendencies embedded in situations, which often depend on mental states such

as expectations and beliefs regarding self-efficacy (Mischel, 1973). In this view, a trait we might consider, such as “ambitious,” is determined by our willingness to pursue our goals, which is in turn affected by self-efficacy beliefs activated in the relevant moment for that goal-directed behavior. People automatically infer traits from the behavior and mental states of others, a phenomenon called spontaneous trait inference (Winter & Uleman, 1984; Todorov & Uleman, 2002; Todorov & Uleman, 2003). The relationship between trait inferences and mental states is born out on a neural level as well. For example, recent work demonstrates that our neural representations of others reflect the mental states we believe they habitually experience (Thornton, Weaverdyck, & Tamir, 2019). Finally, we use our schemas of individuals to infer their mental states in a given situation (Higgins, Rholes, & Jones, 1977). Thus, when we mentalize, we make automatic judgments about others’ transient mental states and enduring traits; these judgments interact and inform one another, and so to understand the perception of short-lived mental states, we must also investigate the perception of stable traits.

There is a rich literature full of two-factor models to explain this essential element of social cognition. Indeed, contemporary psychology has seen the proliferation of two-dimensional models of self, interpersonal, and intergroup perception in a variety of literatures including social roles (Eagly & Steffen, 1984), stereotyping (Fiske, Cuddy, Glick, & Xu, 2002a; Phalet & Poppe, 1997), dehumanization (Haslam, 2006), and mind perception (H. M. Gray, Gray, & Wegner, 2007). These models provide the foundation for the three-dimensional and dynamic approach we describe in this chapter. To provide the context for how and why a three-factor, dynamic model expands beyond previous work, we briefly review various two-factor models.

The two dimensions of these models are referred to by different names in each literature, including self-profitable-other profitable traits (Peters, 1992), agency-communality (Eagly &

Steffen, 1984; or agency-communion (Abele & Wojciszke, 2007), competence-warmth (Fiske, Cuddy, Glick, & Xu, 2002b), agency-experience (H. M. Gray et al., 2007), human uniqueness-human nature (Haslam, 2006), power-benevolence (Leach, Ellemers, & Barreto, 2007), and competence-morality (Phalet & Poppe, 1997; Wojciszke, 1994). Despite differences in nomenclature, there is wide agreement that the core of one dimension reflects social/relational potential, and the other, competence/capacity (either intellectual or physical). While the underlying concepts are largely similar, one way in which the models *do* clearly differ is the way the dimensions incorporate or produce evaluations of morality.

The Stereotype Content Model (Fiske et al., 2002b) suggests that the unique dimensions correspond to the two requisite questions posed by the social world: (1) What are others' intentions?; and (2) What is their capacity to pursue those intentions? (Fiske et al., 2002b). The first question calls attention to, and describes appraisal of, the core dimension *warmth*; the second question, to the core dimension *competence*. In this model, evaluations of morality and sociability are characterized as sub-components of the overarching warmth dimension. In a related conceptualization, Leach (2006) labels a first dimension *power* that encapsulates judgments of competence, strength, prestige, and activity; whereas the second dimension labeled *benevolence* encapsulates judgments of sociality, morality, cooperation, and compatibility. Importantly, both of these approaches treat traits associated with sociability and morality as pro-social and fold them into a single dimension (i.e. *warmth* or *benevolence*) (Leach et al., 2007; see also Fiske et al., 2002; Fiske, 2018).¹

¹ Still other models emphasize a *socio-relational* dimension and a *competence* or achievement dimension (Wojciszke, 2005; Ybarra, Chan, & Park, 2001), or an *intended goal* dimension and *goal attainment* dimension (Phalet & Poppe, 1997). In the domain of face perception, trustworthiness and dominance are seen as the two core dimensions (see Todorov, 2008), though note that a) other work using stimuli that vary more substantially on age obtains a third dimension, related to youthfulness/attractiveness (Vernon et al., 2014), and further, b) more recent

The literature on dehumanization and mind perception also relies on two factor models. However, models in this literature differ in fundamental ways from the two-factor models previously described. In modeling humanness, Haslam (2006) proposes two distinct sets of attributes: (1) characteristics that are *uniquely human*, and (2) characteristics that are aspects of *human nature*. Uniquely human characteristics includes refinement, civility, rationality, cultivation, and moral sensibility. In contrast, human nature includes emotional responsiveness, warmth, cognitive openness, agency, and depth. In this model, morality is found on the first dimension, whereas agency and warmth are found together on the second. The denial of each type of humanness results in a corresponding form of dehumanization—denying human uniqueness results in an implicit vertical comparison that the individual is sub-human, whereas denying human nature results in a horizontal separation that the individual is non-human (Haslam, 2006). Despite locating moral sensibility on a separate dimension from both agency and warmth, Haslam (2012) suggests that the moral status of diverse targets co-varies with the dimensions of humanness (Bastian, Laham, Wilson, Haslam, and Koval, 2011). In this way, the morality of targets appears to be conceived as an emergent property of the two dimensions of the model as opposed to being contained within one of them.

The two-factor model that focuses most explicitly on morality, however, argues that the two dimensions of mind perception—*agency* and *experience*—form the essence of moral judgment (Gray, Young, & Waytz, 2012). In this model, agency, or the perceived capacity to intend and act is orthogonal to experience, or the perceived capacity for sensations and feelings

work suggests that rather than reflecting fixed trait dimensions, social face evaluation is best captured by a dynamic and contextually sensitive framework (Stolier, Hehman, & Freeman, 2018)

(H. M. Gray et al., 2007). In specifying the structure of mind perception, the authors tested 13 characters (types of living humans, non-human animals, a dead woman, God, and a sociable robot) against 18 capacities (H.M. Gray et al., 2007). Factor analysis revealed two dimensions: the *agency* dimension that included the capacities for planning, communication, thought, self-control, emotion recognition, memory, *and* morality; and the experience dimension that included the capacities for hunger, fear, pain, pleasure, rage, desire, personality, consciousness, pride, embarrassment, and joy (H. M. Gray et al., 2007). Unlike most of the models previously described, the two-factor model of mind perception locates the capacity for morality as a sub-component of agency.

The relationships among agency, experience, and morality are complex because the authors also make links between perceiving minds along the agency and experience dimensions, and Dyadic Morality (Gray, Young, & Waytz, 2012), or the attribution of moral rights and moral responsibilities to those minds. Specifically, perceived agency qualifies an entity as capable of doing good or evil, whereas perceived experience qualifies an entity as benefiting from good or suffering from evil (Gray et al., 2012). According to this model, moral judgment depends on a dyadic template of two minds: a moral agent and a moral patient (i.e. the action of the moral agent and the resultant suffering or salvation of the moral patient) (Gray et al., 2012). In other words, perceptions of moral capacity and other competence-related capacities produce perceptions of agency, perceived agency triggers the ascription *moral agent*, and the ascription *moral agent* disqualifies the ascription *moral patient*. At once, targets' perceived capacity for morality informs perceptions of their agency, and simultaneously informs whether they're perceived as good versus bad, or whether their actions are perceived as right versus wrong.

In the next section, we review evidence suggesting that morality is not only a critical and separable dimension of social cognition, but that it may even be the primary dimension. Decades of research have identified the centrality of the warmth and competence dimensions (e.g., Rosenberg et al., 1968), and suggest the universality of these dimensions across cultures (Fiske, Cuddy, & Glick, 2007). Moreover, even in two-factor models, one dimension often takes primacy over the other. For instance, the functionalist perspective invokes evolutionary reasons for the primacy of warmth over competence in social perception. Survival in the social world requires that in encounters with others, individuals must immediately determine whether the target has beneficial or harmful intentions, and only later, whether the target has the ability to enact his or her intentions (Fiske et al., 2007). Along similar lines, goal-oriented approaches to person perception suggest that approach-avoidance underlies impression formation and is more directly based on appraisals of warmth than competence (Wojciszke, Bazinska, & Jaworski, 1998; Wojciszke & Abele, 2008). However, when morality is folded into sociability (or within agency), its independence as a third dimension of social perception, and perhaps the *primary* dimension of social perception, remains obscured.

Morality: The third dimension of social cognition

Are moral evaluations different from other types of evaluations? Is judging a target to be moral different than judging a target to be merely an effective, competent, or sociable one? Aristotle saw morality as the most important basis on which to form positive evaluations, because competence and sociability could only be virtuous, sincere, and trustworthy if expressed through a moral character (MacIntyre, 1984). Aristotle's perspective suggests that morality judgments can color the perception of every other characteristic—consider that immoral competence is dangerous and immoral sociability is disingenuous (Leach et al., 2007; Osgood et

al., 1957). Accordingly, it may be impossible to form *meaningful* evaluations of competence or sociability without first determining whether the target is a moral or immoral actor (see Goodwin et al., 2014; Landy et al., 2016). In this section, we review findings relevant to the moral judgment of targets (and actions), demonstrating that moral judgments appear to be faster and more extreme, privileged, sticky, and affect-laden than non-moral ones, as well as able to shape non-moral ones. We present evidence that moral evaluations are different from other types of evaluation, and in ways that motivate our belief that they may serve a primary role in social cognition.

As suggested in the previous section, morality and sociability information have often been combined into one superordinate dimension, which is often referred to as warmth (e.g., Fiske, Cuddy, & Glick, 2007). However, work over the past decade has begun to tease these dimensions apart, making a distinction between traits associated with morality (e.g., like honesty, sincerity, and trustworthiness (Brambilla, Rusconi, Sacchi, & Cherubini, 2011), as well as fairness, loyalty, courage, etc. (Goodwin, Piazza, & Rozin, 2014), versus traits more generally reflective of sociability (e.g., like likability, warmth, and friendliness (Brambilla et al. 2011), as well as extroversion, agreeableness, enthusiasm, etc. (Goodwin et al., 2014).

Taken together, this work strongly suggests that morality can be dissociated from sociability. For example, a three-factor measurement model that included morality alongside competence and sociability accounted for participants' evaluations of groups better than more parsimonious alternatives (Leach et al., 2007). Specifically, morality was more important to individuals' positive evaluations of their in-group than either competence or sociability. Only morality affected participants' levels of pride in, or social distancing from, their in-group (Leach et al., 2007). Subsequent research has both replicated the primacy of morality in the context of

group impression formation (Brambilla, Rusconi, Sacchi, Cherubini, & Yzerbyt, 2012), and extended this observation to information processing at the level of individual targets.

Specifically, when forming global impressions of other individuals, participants focused more on traits related to morality (like “trustworthy”), versus traits related to either sociability (like “likable”) or competence (like “intelligent,” Brambilla et al., 2011). In addition, morality dominates over sociability and competence information when *updating* impressions as well (Brambilla, Carraro, Castelli, & Sacchi, 2019).

Further work has observed the primacy of morality in person perception, and its dissociability from warmth (e.g., sociability). For example, across seven studies, Goodwin and colleagues (2014) confirmed the domains of morality and warmth are perceived as being distinct from one another, and further, that the moral character associated with a given trait was a better predictor of judgments of its relevance to one’s identity, its desirability, and its controllability as compared to warmth. Further, information related to morality was a better predictor of overall, global impressions of targets than information related to warmth (Goodwin et al., 2014).

Echoing the Aristotelian perspective noted at the beginning of this section, even more recent work indicates that the primacy of morality is such that positive evaluations of both sociability and competence hinge on the presence or absence of moral character (Landy, Piazza, & Goodwin, 2016). For example, sociable and competent targets were only rated positively if they were also moral², whereas moral targets were always rated positively, independent of any other trait information. Taken together, this work provides the basis for the Morality Dominance Hypothesis. While other work has demonstrated that preferences for moral versus immoral traits

² Interestingly, work from the same authors (Piazza, Goodwin, Rozin, & Royzman, 2014) suggests that morality itself may be further parcellated into core goodness traits that amplify moral goodness unconditionally versus value commitment traits (e.g., like “dedicated”) that can amplify moral goodness *or* moral badness conditionally.

may be dynamic and sensitive to context (Melnikoff & Bailey, 2018), morality information drives impression formation, and moral traits are viewed positively—at least to the extent that a perceiver views themselves as being moral (Landy, Piazza, & Goodwin, 2018).

Morality is also separate from and primary to competence and sociability (Brambilla et al., 2011, 2012, 2019; Goodwin et al., 2014; Landy et al., 2016)). Work testing the dominance of moral appraisals over competence appraisals demonstrated the chronic accessibility of morality traits as compared to competence traits, as well as that ascriptions of moral traits better predicted global impressions of targets than the ascription of competence traits (Wojciszke, Bazinska, & Jaworski, 1999). In a related study, the authors tested how both positive and negative morality and competence information was integrated to form impressions of targets. The authors predicted that morality information would have a greater than additive effect on global impressions of targets. The results showed that the evaluative meaning of moral information was independent of competence information (i.e. the moral information retained its direct effect on impression formation irrespective of the target's competence), whereas the opposite was found for competence information. The effect of competence information on global impressions was shaped by the positivity versus negativity of moral information. In other words, morality information provided the necessary context for evaluating whether an individual's competence positively or negatively impacts a global impression of the target, but the same was not true for competence information (Wojciszke et al., 1998), which suggests that moral information is primary.

Research on the how individuals recall information about targets, as well as how individuals revise their prior impressions of targets when they receive more information, also underscores the primacy of morality. For example, researchers explored how participants would

form an impression when the behavioral information (related to both morality and competence) at Time 2 contradicted the behavioral information they received at Time 1 (also related to both morality and competence, but opposite in valence from Time 1; Ybarra, 2011). The results revealed that the amount of change in impressions of the target was greater for the moral domain than the competence domain (when negative behavioral information followed positive behavioral information). These results highlight the greater influence of negative moral information over negative competence information on impression revision processes, when the initial evaluation is positive and more susceptible to adaptation (Ybarra, 2001). Moreover, participants recalled morality information better than either intelligence or neutral information (De Bruin & Van Lange, 1999). We recognize that alone these studies are consistent with a two-factor model (in which morality and sociability are not distinct). But evidence that morality judgments trump sociability judgments and competence judgments suggests that it is weighted most heavily in impression formation.

There are also multiple theoretical accounts suggesting differential processing of moral versus competence information (see Brambilla et al., 2011). One possibility is that perceivers have a lay theory that everyone can behave in moral ways, but only immoral individuals behave in immoral ways. In other words, moral behaviors are less informative about traits, because they may not only reveal underlying characteristics but also reflect situational constraints on behavior. Conversely, perceivers have a lay theory that only competent individuals are capable of great successes, whereas everyone can behave incompetently at one time or another due to situational constraints. Thus, incompetent behaviors are less informative about traits in this domain, whereas competent behaviors are more diagnostic (i.e., they tell you more about the person you are perceiving). These differential expectations about trait-behavior relations produce a

negativity bias in the moral domain because immoral behaviors are more informative, and a positivity bias in the competence domain because competent behaviors are more informative.

A final reason to consider morality as a primary dimension in person perception is that moral evaluations are typically affect-laden (Haidt, 2001), and morality is salient in visual perception (Gantman & Van Bavel, 2015). Affectively-valenced evaluations are made particularly rapidly (Zajonc, 1980) and moral intuitions are especially affect-laden (Haidt, Koller, & Dias, 1993), suggesting that morally-laden rapid inferences of others may occur more quickly than inferences about sociability and competence (though see Fiske, Cuddy, & Glick, 2007). Specifically, within the domain of impression formation, predictions of a target's future behavior in the moral domain were based in part on the perceiver's affective responses to the target, whereas predictions in the competence domain were based only on ascriptions of a relevant personality trait (Brycz & Wojciszke, 1992). In the domain of perception, morally relevant words are more likely to be seen than matched non-moral words (Gantman & Van Bavel, 2014). Together, this evidence is suggestive though not conclusive that morally relevant information may be processed prior to competence and sociability judgments. Empirical evidence demonstrating that moral evaluations occur earlier than other person perception judgments is a promising avenue for future research.

Discussion

We have sought to review evidence that morality should be added to two-factor models of person perception, creating a three-factor model. In this view, moral information exerts a powerful influence on social cognition and influences evaluations within the other dimensions, like sociability and competence. This approach helps account for a wide literature on impression formation and mental state attribution while raising exciting possibilities for future research. In

the final section, we highlight gaps in the current literature and propose some potential avenues for future work in the domain on morality and impression formation.

Moral character judgments may be primary to other types of moral judgments, such as blame and praise, which fundamentally depend on the mental states of the agents. For instance, blame is especially sensitive to intent (Malle, Guglielmo, & Monroe, 2014). Some psychologists have begun to advance a “person-based” as opposed to “act-based” theory of moral blame. They have demonstrated that moral blame can be disproportionate to the actual harm caused by an agent and that relatively harmless acts can receive harsh moral judgments (Pizarro & Tannenbaum, 2011). From this perspective, psychological theories of moral judgment are incomplete because they disregard the primacy of moral character evaluations (e.g. underlying traits, dispositions, and character) to assigning blame and praise, and instead narrowly focus on the local features of the act and agent (e.g., whether the action violates a rule or whether the agent’s mental state at the time of the action allowed for alternative actions, or whether the act caused harm) (Pizarro & Tannenbaum, 2011). For instance, ascriptions of blame to an actor for a car accident were significantly heightened when the actor’s underlying motive for speeding was socially undesirable (to hide a vial of cocaine) versus socially desirable (to hide an anniversary present for his parents; Alicke, 1992). Likewise, rashness leading to an immoral decision was seen as reflecting the wrongdoer’s immoral underlying character, which intensified blame (Critcher, Inbar, & Pizarro, 2013). More work should examine how evaluations of a target’s morality might shape perceptions of the target’s competence or agency.

Moral judgments of an action can also impact whether that action is perceived as having been done intentionally (also known as the Side Effect Effect; Knobe, 2003). In this work, participants learn about a Vice President of a company who approaches the Chairman of the

Board with a project that will increase profits but, as a side effect, will either harm or help the environment. The Chairman of the Board says that he only cares about increasing profits, not the environment. People perceive the harm, but not the help to the environment, as intentional (Knobe, 2003). This underscores that moral character judgments affect perceptions of the agent's mental state. Yet little work has examined how these mental state attributions are allocated across multiple moral agents. For instance, how is moral blame allocated to the Vice President as compared to the Chairman of the Board? And how do other mental state inferences, about deliberation vs. implementation, play in these moral judgments? Given that many moral actions are distributed across multiple moral agents, this will be a fruitful area for future research.

As well, very little work has examined how these judgments unfold over time. Like other evaluations, impressions are likely formed through a series of cycles that occur over time with evaluations being updated and adjusted due to contextual and motivational information (Van Bavel et al., 2012). As such, the final judgment of the morality, competence and sociability of a target are an emergent property of multiple processes unfolding over time (Cunningham & Zelazo, 2007; Cunningham, Zelazo, Packer, & Van Bavel, 2007; Ferguson & Wojnowicz, 2011; Freeman & Ambady, 2011; Scherer, 2009; Van Bavel, Xiao, & Cunningham, 2012). It is unknown how appraisals in each domain might influence appraisal in the others over time. For instance, do moral judgments emerge first in time? If so, this would provide yet another form of moral precedence and might constrain subsequent competence and sociability judgments.

This dynamic approach to social cognition also assumes that context, motivation, and goals will shape attention to, appraisal of, and affective reactions induced by behavioral information about targets, because information processing in general is highly flexible and dependent on both cognitive and motivational goals (Hilton & Darley, 1991; Wojciszke, 2005).

Simply put, goals guide information selection. Thus, perceivers appear to be highly interested in a target's morality in the absence of a particular goal. But, when a domain specific goal is made clear, the task domain instead appears to drive the perceiver's search for corresponding information about the target (Wojciszke, 2005). For instance, when goals are related to competence (e.g. in hiring decisions), perceivers' attention is more directed to competence information. However, when goals are focused on moral judgment (e.g., dating decisions), perceivers' attention is more focused on moral information (see Everett, Pizarro, & Crockett, 2016). Moreover, the value placed on moral vs. immoral traits likely depends on their motivational relevance (Melnikoff & Bailey, 2018). Beyond goals, individual differences are also associated with chronically accessible constructs that perceivers use when evaluating targets and likely play a role in selective impression formation (Wojciszke et al., 1998). Finally, identical actions can be construed in both moral and competence terms depending on the distinct features of the action to which the perceiver is attending (Wojciszke, 1994), and this has important consequences for subsequent evaluations (Van Bavel et al., 2012).

Conclusion

Social cognition is shaped by the *interplay* of three distinct dimensions of morality, sociability, and competence, but moral states and traits may be more relevant to inferring whether a target represents a threat or opportunity than either sociability or competence. Inferences about morality exert their primacy, either by overpowering or by shaping judgments of sociability and competence. It may also be the case that the primacy of morality affects mental state ascriptions. Fully understanding the nature of impression formation, person perception and social cognition is severely impoverished without understanding the power of morality.

References

- Abele, A.E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*(5), 751-763.
- Alicke, M.D. (1992). Culpable causation. *Journal of Personality and Social Psychology, 63*, 368-378.
- Anderson, N.H. (1981). *Foundations of information integration theory*. New York: Academic Press.
- Anderson, C. A., & Sedikides, C. (1991). Thinking about people: Contributions of a typological alternative to associationistic and dimensional models of person perception. *Journal of Personality and Social Psychology, 60*(2), 203-217.
- Bargh, J. A., & Pratto, F. (1986). Individual construct accessibility and perceptual selection. *Journal of Experimental Social Psychology, 22*(4), 293-311.
- Bastian, B., Laham, S. M., Wilson, S., Haslam, N., & Koval, P. (2011). Blaming, praising, and protecting our humanity: The implications of everyday dehumanization for judgments of moral status. *British Journal of Social Psychology, 50*(3), 469-483.
- Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology, 41*(2), 135-143.
- Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology, 51*(1), 149-166.

- Brambilla, M., Carraro, L., Castelli, L., & Sacchi, S. (2019). Changing impressions: Moral character dominates impression updating. *Journal of Experimental Social Psychology, 82*, 64-73.
- Brycz, H., & Wojciszke, B. (1992). Personality impressions on ability and morality trait dimensions. *Polish Psychological Bulletin, 23*(3) 223-236.
- Carnes, N., & Janoff-Bulman, R. (2012). Harm, help, and the nature of (im)moral (in)action. *Psychological Inquiry, 23*(2), 137-142.
- Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. Guilford Press.
- Cikara, M., Farnsworth, R. A., Harris, L. T., & Fiske, S. T. (2010). On the wrong side of the trolley track: Neural correlates of relative social valuation. *Social Cognitive and Affective Neuroscience, 5*(4), 404-413.
- Critcher, C. R., Inbar, Y., & Pizarro, D. A. (2013). How quick decisions illuminate moral character. *Social Psychological and Personality Science, 4*(3), 308-315.
- Cunningham, W.A., & Johnson, M.K. (2007). Attitudes and evaluation: Toward a component process framework. In E. Harmon-Jones & P. Winkielman (Eds.), *Social Neuroscience: Integrating Biological and Psychological Explanations Social Behavior* (pp. 227-245). New York: The Guilford Press.
- Cunningham, W.A., Van Bavel, J.J., & Johnson, I.R. (2008) Affective flexibility: Evaluative processing goals shape amygdala activity. *Psychological Science, 19*, 152-160.
- Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences, 11*(3), 97-104.
- Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative

- reprocessing model: A multilevel framework for attitudes and evaluation. *Social Cognition*, 25(5), 736-760.
- Darley, J. M. (1991). The effects of interaction goals on person perception. *Advances in Experimental Social Psychology*, 24, 236-67.
- De Bruin, E. N., & Van Lange, P. A. (1999). Impression formation and cooperative behavior. *European Journal of Social Psychology*, 29(2-3), 305-328.
- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46, 735.
- Everett, J. A., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, 145, 772-787.
- Fiske, A. P. (2012). Metarelational models: Configurations of social relationships. *European Journal of Social Psychology*, 42, 2-18.
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11, 77-83.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878-902.
- Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, 27(2), 67-73.
- Ferguson, M.J., & Wojnowicz, M.T. (2011). The when and how of evaluative readiness: A social cognitive neuroscience perspective. *Social and Personality Psychology Compass*, 5, 1018-1038.
- Freeman, J.B., & Ambady, N. (2011). A dynamic interactive theory of person construal.

- Psychological Review*, 118, 247-279.
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106(1), 148-168.
- Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, 24(1), 38-44.
- Gray, H.M., Gray, K., & Wegner, D.M. (2007). Dimensions of mind perception. *Science*, 315(5812), 619-619.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101-124.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814-834.
- Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog?. *Journal of Personality and Social Psychology*, 65(4), 613-628.
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, 10, 252-264.
- Haslam, N. (2012). Morality, mind, and humanness. *Psychological Inquiry*, 23(2), 172-174.
- Higgins, E. T., King, G. A., & Mavin, G. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, 43(1), 35-47.
- Higgins, E. Rholes, W.S., & Jones, C.R. (1977). Category Accessibility and Impression Formation. *Journal of Experimental Social Psychology*, 13, 141-154.

- Knobe, J. (2010). Person as Scientist, Person as Moralist. *Behavioral and Brain Sciences*, 33, 315-329
- Landy, J. F., Piazza, J., & Goodwin, G. P. (2016). When it's bad to be friendly and smart: The desirability of sociability and competence depends on morality. *Personality and Social Psychology Bulletin*, 42(9), 1272-1290.
- Landy, J. F., Piazza, J., & Goodwin, G. P. (2018). Morality traits still dominate in forming impressions of others. *Proceedings of the National Academy of Sciences*, 115(25), E5636-E5636.
- Leach, C.W. (2006). *The meaning of prejudice*. Unpublished manuscript, University of Sussex, Brighton, England.
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93(2), 234-249.
- Liu, B. S., & Ditto, P. H. (2013). What dilemma? Moral evaluation shapes factual belief. *Social Psychological and Personality Science*, 4(3), 316-323.
- MacIntyre, A. C. (1984). *After virtue* (Vol. 211). Notre Dame: University of Notre Dame Press.
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, 25, 147-186.
- Melnikoff, D. E., & Bailey, A. H. (2018). Preferences for moral vs. immoral traits in others are conditional. *Proceedings of the National Academy of Sciences*, 115, E592-E600.
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, 80, 252-283.
- Osgood, C. E., & Suci, G. J. (1957). & Tannenbaum, PH (1957). The measurement of meaning.

Urbana: University of Illinois Press.

- Peeters, G. (1992). Evaluative meanings of adjectives in vitro and in context: Some theoretical implications and practical consequences of positive negative asymmetry and behavioral-adaptive concepts of evaluations. *Psychologica Belgica*, 32, 211-231.
- Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, 1(1), 33-60.
- Phalet, K., & Poppe, E. (1997). Competence and morality dimensions of national and ethnic stereotypes: A study in six eastern-European countries. *European Journal of Social Psychology*, 27, 703-723.
- Pizarro, D. A., & Tannenbaum, D. (2011). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. *The social psychology of morality: Exploring the causes of good and evil*, 91-108.
- Pizarro, D., Tannenbaum, D., & Uhlmann, E. (2012). Mindless, Harmless, and Blameworthy. *Psychological Inquiry*, 23(2), 185-188.
- Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, 118(1), 57-75.
- Rai, T.S., & Fiske, A.P. (2012). Beyond harm, intention, and dyads: Relationship regulation, virtuous violence, and metarelational morality. *Psychological Inquiry*, 23(2), 189-193.
- Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles:(Re) introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review*, 1(1), 26-53.
- Reeder, G. D. (1985). Implicit relations between dispositions and behaviors: Effects on

- dispositional attribution. *Attribution: Basic issues and applications*, 87-116.
- Rosenberg, S., Nelson, C., & Vivekananthan, P.S. (1968). A multi-dimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, 9, 283-294.
- Scherer, K.R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, 23, 1307-1351.
- Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, 52(4), 689-699.
- Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. 96(1), 58-83.
- Stolier, R. M., Hehman, E., & Freeman, J. B. (2018). A dynamic structure of social trait space. *Trends in Cognitive Sciences*, 22(3), 197-200.
- Thagard, P. (2004). *Coherence in thought and action*. Boston: MIT Press.
- Tamir, D.I., Thornton, M.A., Contreras, J.M., Mitchell, J.P. (2016). Neural evidence that three dimensions organize mental state representation: rationality, social impact, and valence. *Proceedings of the National Academy of Sciences*, 113(1), 194-199.
- Thornton, M. A., & Mitchell, J. P. (2017). Theories of person perception predict patterns of neural activity during mentalizing. *Cerebral Cortex*, 28(10), 3505-3520.
- Thornton, M. A., Weaverdyck, M. E., & Tamir, D. I. (2019). The brain represents people as the mental states they habitually experience. *Nature Communications*, 10, 2291.
- Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, 83(5), 1051.

- Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*, 39(6), 549-562.
- Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/ avoidance behaviors. In A. Kingstone & M. Miller (Eds.), *The Year in Cognitive Neuroscience 2008, Annals of the New York Academy of Sciences*, 1124, 208-224.
- Trope, Y., & Thompson, E. P. (1997). Looking for truth in all the wrong places? Asymmetric search of individuating information about stereotyped group members. *Journal of Personality and Social Psychology*, 73(2), 229-241.
- Van Bavel, J.J., Packer, D.J., Haas, I.J., & Cunningham, W.A. (2012). The importance of moral construal: Moral versus non-moral construal elicits faster, more extreme, universal evaluations of the same actions. *PlosOne*, 7(11): e48693.
- Van Bavel, J.J., Xiao, Y.J., & Cunningham, W.A., (2012). Evaluation is a dynamic process: moving beyond dual system models. *Social and Personality Psychology Compass*, 6(6), 438-454.
- Vernon, R. J., Sutherland, C. A., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences*, 111(32), E3353-E3361.
- White, G.M. (1980). Conceptual universals in interpersonal language. *American Anthropologist*, 82, 759-781.
- Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, 47(2), 237-252.

- Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology*, *67*(2), 222-232.
- Wojciszke, B. (2005). Morality and competence in person-and self-perception. *European Review of Social Psychology*, *16*(1), 155-188.
- Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, *24*(12), 1251-1263.
- Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology*, *38*(7), 1139-1147.
- Wyer, R. S., & Srull, T. K. (1986). Human cognition in its social context. *Psychological Review*, *93*(3), 322-359.
- Ybarra, O. (2001). When first impressions don't last: The role of isolation and adaptation processes in the revision of evaluative impressions. *Social Cognition*, *19*(5), 491-520.
- Ybarra, O., Chan, E., & Park, D. (2001). Young and old adults' concerns about morality and competence. *Motivation and Emotion*, *25*(2), 85-100.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151-175.