Title
An Inquiry into the Nature and Causes of Choosing Treatment

Author
Martinez, Andres Golden

Publication Date
2012

Peer reviewed|Thesis/dissertation

An Inquiry into the Nature and Causes of Choosing Treatment

By

Andres Golden Martinez

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Psychology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Rodolfo Mendoza-Denton, Chair
Professor Stephen P. Hinshaw
Professor Robert J. MacCoun

Fall 2012

Abstract

An Inquiry into the Nature and Causes of Choosing Treatment

by

Andres Golden Martinez

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Rodolfo Mendoza-Denton, Chair

Despite the existence of many effective treatments for mental health conditions, services are underused. This treatment gap has resulted in substantial psychological, social, and economic costs (Wang, Simon, & Kessler, 2003). Extant research argues that the stigma pervading mental health categories is a significant barrier to seeking needed care (Corrigan, 2004; Link & Phelan, 2006). The current investigation elucidates the psychological processes underlying people's willingness, and reluctance, to seek treatment. Across six studies, employing observational/correlational and experimental designs, evidence converged on a specific psychological process model. Attributing human qualities to (or humanizing) mental illness evokes compassion toward such individuals, which in turn increases perceivers' own willingness to seek treatment, should such care become needed. These results build on previous research—which has examined the differential attribution of humanity in interpersonal and intergroup contexts—by showing how humanizing a stigmatized social group may influence perceivers' own healthcare decision-making. Implications for (de)humanization, emotion, stigma reduction, and promoting wellness are discussed.

Keywords: (de)humanization, stigma, compassion, treatment-seeking

January 8th, 2011 marks the occurrence of a national tragedy. On this fateful day a lone gunman, Jared Lee Loughner, opened fire on a crowd of people, killing six individuals, and severely wounded a U.S. congressional representative. As the nation expressed shock and grief, grappling with what could motivate someone to commit such an offense, it became clear that prior to Jared's deplorable act, he was not psychologically well. Anteceding the tragedy by months were episodes in which Jared displayed erratic, and even bizarre, behavior. Despite overwhelming evidence of his fragile and precarious state of mind prior to the shooting, neither Jared, nor anyone around him, sought help from a mental health professional.

This incident, although extreme and atypical, motivates the question at the heart of the present inquiry: why are so many people who struggle with mental health issues reluctant to seek help from mental health professionals? In the midst of an impressive array of effective evidence-based treatments—both psychotherapeutic and psychopharmacologic—many people avoid psychological services. Estimates suggest that although approximately twenty eight percent of American adults have a diagnosable mental illness annually, only eight percent of American adults seek treatment for mental health conditions (see Marcotte, & Markowitz, 2009). The aim of the present investigation is to shed light on this mental health *treatment gap* (Kohn, Saxena, Levav, & Saraceno, 2004), the disparity regarding those people who would benefit from treatment versus those who actually receive it.

The Treatment Gap

Understanding the treatment gap has important psychological and social consequences. Consider, for example, major depression, a mental health condition characterized by feelings of worthlessness, the inability to experience pleasure, as well as depressed mood (American Psychiatric Association, 2000). Not only does this syndrome affect 16.6 percent of Americans in their lifetime (Kessler, Berglund, Demler, Jin, & Walters, 2005), but 33 billion dollars are lost annually because of reductions in work productivity caused by this disorder (Wang et al., 2003). Furthermore, it is estimated that more than 30,000 Americans take their lives each year (see Hinshaw, 2007), a state of affairs that could be alleviated by greater access to and willingness to seek treatment.

Gaining insight into what motivates people to seek, and avoid, mental health services has additional implications for healthcare. First, for many psychological disorders, there is strong evidence that delaying treatment results in a more problematic illness course, rendering disorders more difficult to successfully manage (Blanco et al., 2008; Post, 1992). Research underscores the importance of early intervention: the sooner people receive treatment, the more favorable their prognosis. Second, it may be cost-effective to treat psychological disorders early, resulting in net healthcare savings (Wang et al., 2003). The United States and other nations struggle with escalating healthcare costs and a strategy of early intervention and preventive care may ease long-term fiscal burdens.

The present inquiry investigates the psychological processes underlying people's unwillingness to participate in mental health treatment, even when it is in their best interest to do so. Adopting a more positive frame, I specifically examine what makes people willing to seek treatment, should they experience psychological distress. A related and important question, which goes beyond the scope of this present inquiry, concerns what makes people stick with treatment once they receive it (that is, treatment adherence). In pursuing the current research question of willingness to seek treatment (hereafter, treatment-seeking), it is crucial to understand why the prospect of mental health treatment is so often met with avoidance, and why

it appears threatening for a great number of people. In short, at issue is why mental health categories (and seeking mental healthcare more generally) are pervaded by stigma.

Social Identity, Stigma, and Mental Health Conditions

Stigma refers to having a social identity that is devalued in a particular context (Crocker, Major, & Steele, 1997; see also Mendoza-Denton, 2008). It may seem inappropriate to refer to mental illness as a social identity, as mental illness is a category belonging to clinical science and medicine: doctors and other mental health professionals treat it, and scientists investigate it. In spite of some assertions that mental illness is an arbitrary social construction (see Szasz, 2007), mental health conditions retain a justified status as biological, or psychobiological, realities (Hinshaw, 2007), although their boundaries are highly contested. Interestingly, the labels of mental illness have a dual function. In addition to being health categories, they also function as social identities that have important social consequences. Just as having melanin in one's skin exerts social effects beyond being a mere physical property of a person, mental illness categories have additional social consequences that extend beyond their associated symptoms. Furthermore, like other social identities, mental illness categories can be situated on dimensions of malleability and valence.

When we think of social identities, gender, race/ethnicity, and sexual orientation are often cognitively accessible. These identities share a property of apparent immutability. However, social identities are not restricted to these often-investigated categories, as they reside on a continuum of malleability. At one antipode are identities that are relatively unchangeable (one's gender, for example, is not easily changed), while at the other extreme are those showing greater fluidity, such as when the identity that we bring to the office dissolves when we come return home from work (Akerlof & Kranton, 2010).

The social identity of mental illness is unusual in its malleability status. Its associated stigma can be willfully adopted or evaded, and also shows a unique trajectory. To become the potential recipient of mental illness stigma, all one must do is make an appointment with a mental health professional. This contrasts with race, nationality, or gender, in that membership in these social categories is largely not of one's choosing. Moreover, refusing to see a clinician can postpone, if not prevent, category membership because one will not receive a diagnostic label. By refusing to acknowledge symptoms, and opting out of any participation in mental healthcare, one can prevent the stigma and discrimination linked to any mental health diagnostic label, a process referred to as *label avoidance* (Ben-Zeev, Corrigan, Britt, & Langford, 2012). However, once category membership is assigned, its effects persist, as the status of *former mental patient* carries an enduring stigma (Link, Cullen, Frank, & Wozniak, 1987). Once a person is in this category, however, mental illnesses take on the seemingly irreversible characteristics of other social identities, such as those stemming from race or gender.

In addition to a range of malleability, social identities also take on a range of valences. Some identities are positively valenced and admired. For example, the social category of middle-class Americans is highly valued among citizens the United States (Fiske, Cuddy, & Glick, 2006). In contrast, other identities are negatively valenced, being the objects of scorn. An example here, for many Americans, would be the social category of people receiving public assistance. Stigma is located on this lower end of the continuum, representing a devalued identity, and this where people bearing mental illness labels reside.

Among the social categories to which one can potentially belong, the negative valence, disesteem, and devaluation associated with mental illness is especially pronounced. When perceivers are confronted with behaviors that sometimes covary with, and are stereotypically

associated with, mental illness (such as homelessness and abusing substances) the medial prefrontal cortex—a brain region associated with thinking about other people—fails to activate (Harris & Fiske, 2006). Furthermore, viewing these targets activates the insula, a neural region associated with feelings of disgust, further suggesting an attribution of subhuman status to these targets. These results argue for a neural signature of dehumanization that is evoked by cues stereotypically associated with mental illness. Furthermore, merely imagining a person bearing a mental illness label (in comparison to a physical illness label) triggers reduced attributions of humanness to the target (Martinez, Piff, Mendoza-Denton, & Hinshaw, 2011, Study 1). In other words, when thinking about a person with a mental illness label, the default attribution may be (on average) dehumanization.

In sum, these recent findings suggest that attributions of humanness, and its denial, may be especially important for understanding the psychological and behavioral processes involved in mental illness stigma. This dimension of social perception that ascribes or denies humanness to others has only recently been subject to systematic investigation. In my research, I examine how (de)humanization may influence self/other representations, induce particular emotional responses (compassion) and eventuate in specific healthcare behaviors. The process of (de)humanization, however, is multifaceted and I turn my attention to this important social cognitive process.

The Continuum and Forms of (De)humanization

Humanness, and its denial, is a critical dimension of social perception (Haslam, Loughnan, Kashima, & Bain, 2008). It signifies the degree to which attributes of humanness are ascribed (or denied) to a human target, social group, or physical object. One pole of this continuum, reflecting the denial of human attributes to a target (dehumanization), has received the bulk of empirical attention. Beginning with work on infrahumanization—the differential attribution of uniquely human emotions to a person or group (Leyens, Demoulin, Vaes, Gaunt, & Paladino, 2007)—a vibrant literature has emerged that has elucidated the various forms that humanization and its denial may take.

Recent theoretical and empirical work demonstrates that there are at least two forms of the denial of humanity. In one influential taxonomy, targets can be likened to animals or androids (Haslam, 2006; Haslam et al., 2008). With animalistic dehumanization, targets can be denied uniquely human emotions that separate humans from animals (such as the capacity to experience admiration, guilt, or nostalgia) or they can be ascribed characteristics of animals (targets can be seen as beastly and uncivilized). Researchers have operationalized this form of dehumanization in a variety of ways, such as differentials in the attributions of uniquely human emotions to ingroups versus outgroups (Demoulin et al., 2004), ascription of words related or animality or humanity (Viki et al., 2006; Zebel, Zimmermann, Viki, & Doosje, B., 2008) or the denial of personality characteristics that lay perceivers uniquely attribute to humans, such as openness and conscientiousness (Costello & Hodson, 2010; Hodson & Costello, 2007). Further, animalistic dehumanization can occur without conscious awareness, as investigators have used reaction time methodologies to measure implicit dehumanization (Loughnan & Haslam, 2007; Paladino et al., 2002).

Whereas some targets can be seen as lacking qualities that make them uniquely human, rendering them animalistic, others targets may be seen to lack human nature (Haslam, 2006). That is, they may lack the qualities of warmth, vitality, and vibrancy that distinguish humans from machines. To illustrate, whereas an artist may seem to possess warmth, spontaneity, and the impulses that constitute human nature, businessmen or women may be seen as robotic, cold, inert, and lacking the human feelings that differentiate them from androids (Loughnan &

Haslam, 2007).[1] Furthermore, when people look inward to perceive themselves, they may attribute greater human nature to themselves versus others (Haslam, Bain, Douge, Lee, & Bastian, 2005). This latter effect may be the result of people's privileged access to their own psychological states.

A recent complementary conceptual orientation argues that the uniquely human and human nature distinction can be mapped on to two distinct qualities of mind: *agency* and *experience* (Gray, Gray, & Wegner, 2007; Waytz, Gray, Epley, & Wegner, 2010). In this mind perception view of humanization, *agency* refers to the capacity to plan, intend, and think, whereas *experience* refers to the capacity to sense, feel, and experience pain. These theorists contend that *agency* closely maps onto the notion of the *uniquely human*, as people preferentially attribute these mental capacities to humans rather than animals, whereas *experience* closely maps onto *human nature*, as these latter mental capacities distinguish those who can feel, versus entities (e.g., androids) that cannot (Waytz et al., 2010).

Recent speculation suggests that there may even be an additional form of dehumanization, distinct from human uniqueness and human nature. This perspective explores how dehumanization can involve likening individuals to inanimate matter. Whereas animals and androids both have a form of agency, in that they exert effects on the world, other entities do not. Consider a piece of trash, which cannot agentically affect the world in any fashion, and is simultaneously highly devalued. Some individuals, such as the homeless, may be dehumanized when perceivers liken them to refuse and thus fail to ascribe them human, even animate, status (see Waytz, Epley, & Cacioppo, 2010). Consequently, perceivers may treat such people with indifference and lack of concern, easily strolling by and ignoring them as they express their needs.

Although (de)humanization is usually about *people* being denied or ascribed humanness, non-human objects and organisms can also be ascribed humanness. For example, we may see our computers and other electronic devices as humans, speaking soothingly to them to coax them into functioning, or yelling at them when they lag or malfunction (see Epley, Waytz, & Cacioppo, 2007). Further, we may also humanize our pets (Epley, Akalis, Waytz, & Cacioppo, 2008), seeing our animal companions (such as our dogs and cats) as possessing attributes typically associated with humans. Humanizing or anthropomorphizing non-human entities may be accentuated by loneliness and may provide comfort from isolation.

In spite of the recent advances in this field, the varieties of humanization may be more diverse and complex than empirical research has acknowledged. Humanness is viewed, almost as a default, as a desirable and positive attribute. However, humanness is fraught with more meanings than either the Enlightenment tradition of civility and reason (in contrast to animals) or the Romantic tradition of vitality (in contrast to androids), a distinction Haslam (2006) eloquently makes. Pioneers in this field, however, acknowledge that our understanding of humanization is in its nascent phases. To further enrich (if not complicate) matters, humanness can also refer to weakness, frailty, and fallibility, or being "*only human*" (Haslam, Loughnan, Reynolds, & Wilson, 2007). Consider a marriage partner succumbing to infidelity. In this case, attributions of humanness to one's partner refer to an instance of transgressing a moral boundary, or failing to live up to social norms, community standards, and even abstract ideals. Here, an attribution of humanness depicts people as weak and at their lowest, not their greatest glory, providing reason to absolve them from moral failures and sins.

This discussion highlights the complexity of the seemingly straightforward concept of humanness. Whereas attributions of humanness and its denial may invoke animality, robotics,

inert matter, anthropomorphism, and moral decay, the present inquiry will focus on a particular form of humanization. Specifically, I will concentrate on the dimension on which lower values denote animality versus (at the other extreme) ascription of Enlightenment notions of humanity. The upper end of this dimension is indexed by concepts and processes denoting reason and complex thought, the capacity to experience complex emotions, and uniquely human characteristics such as openness.

Examining mental health stigma by targeting this specific form of humanization has a compelling rationale. Historically, representations of people with mental illness have often depicted them in humiliating ways, and such individuals were often subject to cruelty and abuse as they were deemed incapable of reason (Hinshaw, 2007). Further, contemporary media images typically depict individuals with mental illness as out of control and violent (Sieff, 2003; Wahl, 1995), unable to act reflectively and deliberatively. In contrast, attributions of the emotionless android (that is, denial of human nature) seem implausible, as many perceivers spontaneously associate mental illness with dangerousness (Ben-Zeev et al., 2012; Link, Phelan, Bresnahan, Stueve, & Pescosolido, 1999; Martinez et al., 2011)—a characteristic that implies an excess, rather than a deficit, of passion. Guided by this reasoning, I investigate humanization linked to the Enlightenment versus animalistic conceptions of humanness, acknowledging the variability in the extent to which people humanize the category of, and individuals with, mental illness.

(De)humanization and Compassion

Humanization may have downstream affective consequences, one of which may be compassion. The emotion of compassion has received extensive theorizing, having roots in the moral philosophies of both Eastern and Western traditions. From the perspective of Eastern philosophy, compassion is viewed as the central moral emotion to cultivate, as proposed by key figures such as the Dalai Lama (see Keltner, 2009 for a discussion). Further, compassion (also referred to as sympathy or fellow feeling) is a central pillar of Western philosophy, receiving its most extensive theoretical elaboration by Adam Smith in the 18th century (Smith, 1759/2002). Writing a century later, Darwin incorporated Smith's ideas into his own biological theorizing, arguing that human communities possessing this emotion (or sympathy, as he called it) would be more likely to survive, as they needed to care for and help each other, especially vulnerable offspring (see Ekman, 2010). Compassion is critical for human survival because it facilitates a care-taking response.

Although compassion has gone by a variety of monikers, include sympathy and empathy (Ekman, 2010; Geotz, Keltner, & Simon-Thomas, 2010), of central importance is compassion's core feature and associated behavioral tendencies. Compassion has been conceptualized as an other-focused emotion that orients the individual toward promoting the well-being, and reducing the suffering, of another person (Geotz et al., 2010; Keltner, Marsh, & Smith, 2010). What facilitates experiencing this concern remains insufficiently explained. I investigate how humanizing may influence compassion, as humanizing may facilitate consideration of another's cognitive and emotional states (Carman, 2009). This may happen by including the perceiver and outgroup within the same category of humans (Sayer, 2005), thereby allowing the perceiver to more easily enter into the specific concerns of the target. Recent research leads credence to this claim. When compassion is experimentally induced (by showing images of human suffering) participants who view these images later perceive themselves as more similar to people from disparate social categories, although targets with mental illness were not included in these studies (Oveis, Horberg, & Keltner, 2010). To the extant that humanizing facilitates an inclusive self-representation that incorporates stigmatized others, compassionate concern toward them may

result. I hypothesize that when compassion toward a stigmatized other is activated by humanization, this compassionate concern may have downstream consequences for the perceiver's own behavior.

(De)Humanization, Compassion, and Treatment-Seeking

The overarching theoretical framework is this inquiry is rooted is the broad claim that emotions mediate the linkage between cognitions and behavior. As Bernard Weiner succinctly stated (see Siegel & Shaughnessy, 1996), "Emotions provide the bridge between thinking and doing" (p.172). In what follows I describe the hypothesized model incorporating the emotional and behavioral consequences of humanizing. The proximate psychological mechanism that underlies a link between humanizing mental illness and perceiver treatment-seeking should be an emotion because emotions are psychological states that motivate action (Frijda, Kuipers, & ter Schure, 1989).

The central hypothesis is that humanizing a target with mental illness evokes compassion toward this target and thereby influences perceiver's own treatment-seeking tendencies, should the perceiver experience psychological difficulties. In this way, humanization triggers the care-taking response associated with compassion, which can become re-directed toward the perceiver, should the perceiver become vulnerable. Humanization may set this process in motion, because the stigmatized other (a person with mental illness) is included in the perceiver's self-representation.

Although there is no direct evidence bearing on this process model, several studies indirectly lend plausibility to this theorizing. Throughout this research, however, emotional mediators were not assessed. Recent research shows a general connection between humanization and prosocial behavior, specifically protecting or helping another. Although in a somewhat remote domain, attributing uniquely human mental states to animals is associated with a reluctance to use them for food (Bastian, Loughnan, Haslam, & Radke, 2012). In this way, humanizing a vulnerable other (even a member of a different species) is associated with a desire to abstain from harming or exploiting them. More relevant to the current research is the relation (among human participants) between humanization of stigmatized others and altruistic acts directed toward them (Cuddy, Rock, & Norton, 2007). In this study, attributing uniquely human emotions to an outgroup (a form of humanization) predicted perceivers' willingness to provide them with help. In other words, humanizing others influenced the general motivation to give aid or provide others with care. Further, and bearing directly on mental health stigma, humanizing an individual diagnosed with a mental illness diminishes perceivers' motivations to socially exclude him (Martinez et al., 2011, Study 2). Alternatively expressed, humanizing a stigmatized target with a mental illness label may augment approach tendencies as well as inclusion. Overall, as perceivers humanize stigmatized individuals, this process influences perceivers' desires to protect, provide aid to, and include them. It is likely that humanization can activate other behaviors that are linked with care and support. It is possible that this care-taking impulse, the hallmark of compassionate responding, can direct inwardly, should the self need care. In this investigation, I examine how humanization of a stigmatized other can eventuate in treatment-seeking (a form of care directed toward the self) while testing for a specific emotional/motivational mediator: compassion.

Overview of the Present Research

The present research investigates how humanization of mental illness can eventuate in increases in perceivers' willingness to seek treatment. I investigate a psychological process model whereby humanization of mental illness influences compassion toward members of this

category and, in turn, increases perceivers' own treatment-seeking. Across studies, different operationalizations of the key constructs are employed to test the robustness and generalizability of the model. The first study is an initial test for a link between humanization of mental illness as a category and treatment-seeking. Whereas Study 1 employs student samples, Study 2 tests whether this linkage generalizes to a nationwide sample. Study 2 also examines my basic theoretical model (humanization → compassion → treatment-seeking) in an observational/correlational context. Study 3 complements the previous study by experimentally inducing humanization and replicating the model. In addition, Study 3 pits my hypothesized psychological process model against rival alternative models.

Studies 4 through 6 decompose, and further probe, the effects investigated in the first 3 studies. Whereas Studies 2 and 3 correlate humanization and compassion by analyzing measured variation, Study 4 manipulates humanizing, in a novel context, to assess the causal impact of humanization on compassion. Study 5 further probes the humanization-compassion link by testing for an underlying inclusive mental representation (inclusion-of-the-outgroup-in-the-self) that humanization may evoke. In this way, Study 5 addresses the underlying mental representations that allow humanizing to evoke compassion. Importantly, investigating inclusion-of-outgroup-in-the-self as a mediator of the humanization → compassion in Study 5 allows building and testing a more nuanced psychological process model. I specifically test humanization → inclusion-of-outgroup-in-the-self → compassion → treatment-seeking. Study 5 also employs an alternative measure of treatment-seeking, one capturing incremental concrete healthcare-seeking behaviors.

Finally, Study 6 investigates the specific determinants of humanizing a person with mental illness by varying information provided about a target's course of treatment. As the informational manipulation in Study 3 combines several elements to elicit the humanization effect, Study 6 dismantles this manipulation to ascertain the critical ingredients. Specifically, Study 6 employs different information regarding the target's treatment status and treatment effectiveness to determine what prompts differential humanization within a mental health context.

Study 1: Humanization of a Stigmatized Target is Associated with Perceiver Treatment-Seeking

I conducted an initial observational/correlational study to determine if humanizing the category of mental illness influences treatment-seeking. The outcome variable measured the degree to which a person would be willing to seek treatment in the event that she or she would experience psychological distress on a future occasion. A multiple regression analysis was employed to control for a variety of demographic variables, including gender, ethnicity, age, political orientation, and socioeconomic status. Studies have found that mental health stigma may vary by ethnicity (Loya, Reddy, & Hinshaw, 2010) as well as gender (see Mehta & Farina, 1997). I therefore adjusted for these potential confounds to ascertain humanization's unique influence on willingness to seek treatment.

Method

Participants

Participants were 129 students at a large public university who received partial course credit for their participation. Five participants were excluded because experimenter error, resulting in a final sample of 124 participants (62.1 percent female). The ethnic composition was 34.7 percent European American, 38.7 percent Asian American, with 26.6 percent stating other

ethnicities. The average age was 20.62 years old (*SD* = 3.30). The mean political ideology (1 = *strongly liberal*, 7 = *strongly conservative*) was 2.79 (*SD* = 1.47). Socioeconomic status was assessed on a 5-point scale (Côté, Gyurak, & Levenson, 2010; Horberg, Oveis, Keltner, & Cohen, 2009) coded from 1 (*lower class)* to 5 (*upper class)*. This measure is strongly related to reports of household income (*r* = .51, Stellar, Manzo, Kraus, & Keltner, in press). The mean for this sample was 3.02 (*SD* = 1.03). Across studies, I employed these demographic measures.

Procedure

Participants were greeted by a female experimenter and entered a private cubicle in which they completed a series of tasks and surveys on a computer. Participants engaged in these tasks alone and did not provide any identifying information. Upon completion of the study, participants were thanked, debriefed, and awarded partial course credit.

Measures

Humanization. To assess humanization, participants used a Likert rating scale (1 = *not at all*, 7 = *totally*) to indicate the extent to which they explicitly associated human words (*person*, *human*, *civilized*, and *citizen)* and animal words (*animalistic*, *wild*, *beastly*, and *untamed*) with mental illness (Martinez et al., 2011; Viki et al., 2006; Zebel et al., 2008). Animal words were reverse-scored and then combined with the human words by standardizing the humanity and animality indices and averaging them. This yielded a composite humanization measure in which lower scores indicate dehumanization and higher scores indicated increased humanization, α = .88.

Treatment-seeking. I measured treatment-seeking by using a scale that assesses willingness to seek care from a mental health professional, conditional on experiencing psychological distress. Specifically, I employed eight items of the help-seeking propensity subscale of the Inventory of Attitudes Toward Seeking Mental Health Services (Mackenzie, Knox, Gekoski, & Macaulay, 2004). Sample items included: *"I would want to get professional help if I were worried or upset for a long period of time"* and *"If I believed I were having a mental breakdown, my first inclination would be to get professional attention."* The scale anchors were 1 = *strongly disagree*, 5 = *strongly agree*, α = .82.

Results and Discussion

A bivarate regression showed that humanization of the category of mental illness predicted treatment-seeking (β = .28, $p < .01$). To test robustness of this humanization effect, a multiple regression entered all demographic covariates as predictors. Specifically, treatment-seeking was regressed on the key explanatory variable (humanization) as well as gender (female = 0, male =1), political ideology, ethnicity (1 = European American, 0 = other), age (coded 18 to 24 =0, 25 and over as 1), and socioeconomic status. The humanization variable reduced in magnitude but remained significant in the presence of these potential confounding variables (β = .19, $p < .05$).[2]

This study provides initial evidence that explicitly associating human attributes to—or humanizing—the category of mental illness has downstream influences on perceivers' treatment-seeking, in the event that he/she should develop mental health symptoms. In addition, humanization showed predictive power in the presence of a variety of potential demographic confounds (including gender, ethnicity, socioeconomic status, age, as well as political ideology). The next study seeks to go beyond this bivariate relation to test a more complex process model.

Study 2: Humanization, Compassion, and Treatment-Seeking

Study 2 tests my psychological process model with an observational/correlational design. Specifically, I test the model in which humanization (of the category of mental illness) predicts compassion toward members of that category, which in turn influences perceivers' own treatment-seeking tendencies (humanization → compassion → treatment-seeking). Stated differently, humanization leads to treatment-seeking via the emotion of compassion felt toward people with mental illness. Study 2 also tests an alternative explanation to the model. It is possible that humanization is substituting for valence and that valence is the true explanation for any findings. Afterall, humanness is often a desirable and positive attribute and it is possible that humanness attributions are indexing nothing more than liking. I provide a test of this alternative account by partialling out ratings of dangerousness from the model. These latter ratings may be construed, at minimum, to proxy for valence.

Method

Participants

Participants were recruited for an online study through a national online retailer's data collection service and received monetary compensation in exchange for their participation. Across online studies, data were screened for repeat responders using computer IP addresses (Kraut et al., 2004) and only US citizens were included. Of the 425 participants recruited, five were excluded because of substantial missing data. To ensure data quality I included a "catch trial" or comprehension check (Paolacci, Chandler, & Ipeirotis, 2010) and a priori excluded participants who incorrectly identified the main content of a passage they had read (twenty four percent of the total sample), resulting in a final sample of 320 participants (56.3 percent female). The ethnic composition was 75 percent European American, 8.1 percent Asian American, with 16.9 percent stating other ethnicities. The average age was 32.85 years old (SD = 12.37). The mean political ideology was 3.37 (SD = 1.70) and the mean socioeconomic status was 2.63 (SD = .83).

Procedure

Participants were invited to participate in the study by clicking on a link that included a variety of measures. Upon entering the study, participants were randomly assigned to read a passage that described varying perspectives on the origins and treatments of mental health conditions. However, since this manipulation did not exert systematic effects on subsequent measures, I collapsed across it. Further, including an indicator variable coding for condition in subsequent analyses did not alter the statistical significance of any of the findings. After participants completed the measures, they were thanked, debriefed, and then they were awarded their compensation.

Measures

Humanization. Participants completed the same humanization measure as in Study 1, except "beastly" was replaced with "beast," and "animalistic" was replaced with "animal," α = .83.

Compassion. Compassion was measured with a scale used in previous research (Batson et al., 1997) in which participants rate the extent to which they felt six emotions, including compassion, toward a target. Ratings were made toward people diagnosed with mental illness and employed a 7-point Likert scale anchored at 1 = *not at all* and 7 = *very much*, α = .94. Although this index has been used to assess what Batson and colleagues (1997) call empathy, it more closely aligns with recent theorizing on compassion as it includes feeling states directed

toward a target, rather than mirroring a target's emotions or cognitions (Goetz et al., 2010; Waytz & Mitchell, 2011).

Dangerousness. I assessed the extent to which participants associated dangerousness with mental illness using a seven-point Likert scale (1 = *not at all*, 7 = *totally*). The specific items were *violent*, *chaotic*, *out-of-control*, and *danger*, $\alpha = .90$.

Treatment-seeking. I measured treatment-seeking with the same measure employed in Study 1, $\alpha = .85$.

Results and Discussion

Replicating Study 1, humanization predicted treatment-seeking, although with reduced magnitude $\beta = .19$, $p < .01$. My process model predicts that humanization evokes compassion, which in turn influences treatment-seeking. I therefore conducted a bootstrapped mediation analysis to assess the viability of this model (Hayes, 2009; Preacher & Hayes, 2004, 2008). Figure 1 graphically depicts the results of this analysis with standardized regression coefficients. Supporting mediation, humanization predicted the mediator, compassion ($\beta = .29$, $p < .001$), and this mediator predicted treatment-seeking when controlling for humanization ($\beta = .38$, $p < .001$). Full mediation was attained, as the humanization/treatment-seeking link was no longer significant when the mediator (compassion) was included in the regression model ($\beta = .08$, *ns*).

If the humanization/treatment-seeking link is indeed mediated by compassion, the bootstrapped estimate of the indirect effect will have a confidence interval that does not include zero.[3] Results indicated a significant indirect effect, as the bootstrapped 95% confidence interval did not include zero (CI: .06, .17). Testifying to the robustness of this model, the indirect effect remained significant when the model was re-conducted with all demographic covariates included (CI: .05, .16). Furthermore, the indirect effect remained significant after entering these same demographic covariates alongside the dangerousness measure (CI: .04, .16).

These results support the process model in which humanization predicts treatment-seeking, via feelings of compassion toward people with mental illness. These results also argue that this effect is independent of valence, as controlling for ratings of dangerousness, a clearly negative rating, did not alter the pattern of the results. Moreover, the effect remained after controlling for a host of demographic confounds. Although these findings may be intriguing, the study was correlational; no variables were manipulated. The next study seeks to directly induce humanization with a manipulation and then determine whether the mediational pattern replicates.

Study 3: Experimentally-induced Humanization Evokes Compassion and, in turn, Treatment-seeking

Whereas Studies 1 and 2 measured variation in chronically accessible associations of humanness to mental illness, Study 3 seeks to directly experimentally induce humanization, thereby experimentally activating the process in which humanization leads to compassion and, in turn, treatment-seeking. Also, Studies 1 and 2 examined humanizing of the category of mental illness (participants were asked how much they associated certain attributes with mental illness). Study 3 builds on these findings by having participants rate a particular target. In so doing Study 3 tests whether the findings of Studies 1 and 2 extend to rating an individual.

Method

Participants

Participants were 92 students at a large public university who received partial course credit for their participation. One participant was excluded due to a computer malfunction and 4

participants were excluded because of experimenter error. Finally, 3 participants were excluded because of excessively fast times on a reaction time measure (more than 10 percent of their trials showed latencies less than 300 milliseconds; Greenwald et al., 2003). This resulted in a final sample of 84 participants (73.8 percent female). The ethnic composition was 27.4 percent European American, 46.4 percent Asian American, with 26.2 percent stating other ethnicities. The average age was 21.10 years old (*SD* = 3.82). The mean political ideology was 2.90 (*SD* = 1.43) and the mean socioeconomic status of the sample was 3.27 (*SD* = .96).

Procedure

Participants were greeted by a female experimenter, escorted into a sound-attenuated cubicle, and randomly assigned to one of two conditions. The entire study was administered on a computer. After completing individual difference measures and filler items, participants experienced either the experimental condition (hereafter, the recovery condition) or the control condition. The recovery condition presented a vignette about a person (hereafter, target) described as being diagnosed with a mental illness, but having experienced a prolonged recovery. This target, Donald, behaves benignly and enjoys an ordinary day with his friend. The vignette (Srull &Wyer, 1979) is told from the perspective of the friend and was modified to eliminate all ambiguously hostile content. The vignette presented a series of benign, and mundane, events that were held constant across conditions.

Participants read:

> "I ran into my old acquaintance Donald the other day. He is diagnosed with a mental illness but he has been successfully treated for as long as I can remember. I decided to go over and visit him, since by coincidence we took our vacations at the same time. We talked for a while, had lunch, and then went out for a ride. We used my car. We went to the park for about an hour and then stopped at a hardware store. I couldn't find what I was looking for, so we left and walked a few blocks to another store. It was getting kind of late, so I took Donald to pick up his car and we agreed to meet again as soon as possible."

The control (or baseline) condition contained identical behavioral information, but no information about diagnostic status or treatment was mentioned.

After participants read the vignette, they were asked to write a few sentences about their impressions of the target. This was done to strengthen the manipulation via elaborative processing. After writing about the target, participants completed the dependent measures as well as a demographic form. After that, participants were thanked, debriefed, and given partial course credit.

Measures

Humanization. Participants completed the same humanization measures as in Study 1, with one difference. For greater conceptual coverage, I supplemented this measure by including an indirect measure of uniquely human personality characteristics ascribed to Donald, a composite of openness and conscientiousness (Costello & Hodson, 2010; Hodson & Costello, 2007), measured with the Ten Item Personality Inventory or TIPI (Gosling, Rentfrow, & Swan, 2003). To create a single humanization index, I first reverse-scored animality and then combined the measures by z-scoring and averaging them (Martinez et al., 2011), with higher scores indicating increased humanization, $\alpha = .74$.

Compassion. Participants completed the same compassion measure as in Study 2, but rated their feelings toward the target, $\alpha = .92$.

Treatment-seeking. Treatment-seeking was measured with the same scale employed in Studies 1 and 2, $\alpha = .90$.

Results and Discussion

The first set of analyses sought to rule out valence of the target as an explanation for any findings. Participants may have associated more positive valence to one target over another, and thus differential valence could confound interpretation of other effects. To examine this possibility, I compared the two conditions on ratings of the target on agreeableness from the TIPI. Across conditions, I compared the aggregated agreeableness index, as well as each item individually, to rule out this alternative explanation. Across all three tests there was no evidence for a differential positive valence interpretation (all *p*s > .10)

The next set of comparisons tested my main hypothesis: whether the recovery condition, relative to the baseline condition, increased humanization of the target as well as increased treatment-seeking. Analyses yielded evidence for both predicted effects. The recovery condition increased humanization (M = .16, SD =.60) relative to the control condition (M = -.15, SD = .66), $F(1,82) = 4.90$, $p < .05$, $d = .49$.[4] Furthermore, the recovery condition increased treatment-seeking (M = 3.51, SD = .96) relative to the control condition (M = 2.84, SD = .90), $F(1,82) = 10.72$, $p < .01$, $d = .72$.[5]

Next, I tested my hypothesized psychological process model. This model specifies that humanizing a target with mental illness increases compassion toward the target, which in turn increases the perceivers' own treatment-seeking tendencies. In the context of this experiment, I tested for the following path: condition → humanization → compassion → treatment-seeking. I also included a direct path from condition to willingness to seek treatment. I assessed model by employing four indices recommended by Kline (2005): Model Chi-Square (χ^2), Comparative Fit Index (*CFI*), Root Mean Square Error of Approximation (*RMSEA*), and Standardized Root Mean Square Residual (*SRMR*). Condition was the exogenous variable (coded 0 = control condition, 1 = recovery condition). The hypothesized model showed good fit across indices: χ^2 (2) = 1.45, $p > .48$, *CFI* = 1.00, *RMSEA* = 0.00 (90 percent confidence interval: 0.00 to .20), *SRMR* = .04. Figure 2 depicts this model with standardized path coefficients.[6]

In sum, within an experimental context, Study 3 replicated the psychological process model in which humanization leads to compassion and, in turn, increases perceivers' own treatment-seeking. Building on Study 2, this process was set in motion by an exogenous manipulation. In addition, Study 3 shows that the humanization effects obtained in Study 2, in which participants associated human attributes to a category, extend to particular targets. Studies 2 and 3 are limited, however, in that they link humanization and compassion by simply analyzing variation. The causal status of humanization in evoking compassion remains untested. The next study experimentally probes this humanization-compassion link.

Study 4: The Causal Impact of Humanization on Compassion

In the next study I tested whether humanization evokes compassion by experimentally manipulating three levels of humanization of a target and then measuring compassion felt toward this target. For the target I used an image of a canine with three different descriptions to evoke differential humanization: low, medium, and high. The decision to use this stimulus was motivated by the aim of eliciting differential levels of humanization, not investigating canines per se. I chose this canine stimulus because research suggests that dogs may especially susceptible to humanization (Epley et al., 2008; see also Kwan & Fiske, 2008). My central prediction is that increases in humanization of a target (relative to a baseline) will evoke heightened feelings of compassion toward the target.

Method

Participants

Participants were recruited for an online study through a national online retailer's data collection service and received monetary compensation in exchange for their participation. Of the 192 participants that were recruited, two were excluded because of substantial missing data and one was excluded for not writing anything for the writing prompt, which served as the "catch trial" or compliance check for this study. Of the remaining 189 participants (52.4 percent female, 1 participant did not state gender), 77.2 percent were European American, 4.8 Asian American, with 18 percent stated other ethnicities The average participant age was 37.22 years old (*SD* = 12.92). The mean political ideology was 3.15 (*SD* = 1.79) and the mean socioeconomic status was 2.48 (*SD* = .89).

Procedure

Participants were randomly assigned to one of three conditions in a between-participants design. In each condition, participants saw a color image of a canine (specifically, a Tri-Color Pembroke Welsh Corgi), were provided with a description of this target, and then were asked to write two sentences regarding their impressions of him. All participants read the following initial description:

> *"This is a picture of a dog named Rex. He is a very energetic dog who loves to play. He enjoys spending time at the local dog park and playing with other dogs. He especially enjoys running long distances to retrieve a ball after his caretaker throws it."*

In the high-humanness condition, suggesting high levels of uniquely human personality traits, participants additionally read:

> *"Rex, however, is particularly curious about the world around him. He spends a great deal of time exploring and investigating his surroundings. In addition, Rex makes great efforts to secure the family home from potential intruders. He spends a lot of time near the front door, seeming concerned about who may enter the family home."*

In contrast, participants in low-humanness condition, suggesting low levels of uniquely human personality traits, read:

> *"Rex, however, is not particularly curious about the world around him. He does not spend much time exploring or investigating his surroundings. In addition, Rex makes little effort to secure the family home from potential intruders. He spends very little time near the front door, seeming unconcerned about who may enter the family home."*

In the baseline condition, participants read the initial description and then immediately proceeded to the writing task. This design yielded three ascending levels of humanness: low-humanness, a baseline, and high-humanness. After completing the writing task, participants rated attributes of the target, as well as their feelings toward the target. After completing a demographics form, participants were thanked and then debriefed.

Measures

Humanization. Humanization was measured with items indexing openness and conscientiousness from the TIPI (Costello & Hodson, 2010; Hodson & Costello, 2007) as in Study 3 and were supplemented with items from the mind attribution scale (Kozak, Marsh, & Wegner, 2006), a measure employed to assess (de)humanization in recent research (Waytz & Epley, 2012). Sample items from the mind attribution scale include *"This person is capable of planned actions"* and *"This person has complex feelings."* These items were rated on a 7-point Likert scale, anchored at 1 = *strongly disagree* and 7 = *strongly agree*. Each item was worded so that it described the canine target, Rex. These humanization indices were aggregated via z-

scores, $\alpha = .82$.

Compassion. Compassion was measured with a three-item composite index used in previous research (Oveis et al., 2010, Studies 2 and 3). Specifically, participants rated the extent that felt "*compassionate*," "*sympathetic*," and "*moved*" on a 7-point Likert scale anchored at "*not al all*" and "*very much*." Participants were asked to indicate the extent to which they experienced these feelings "*toward Rex*." This index was internally consistent, $\alpha = .80$.

Results and Discussion

To ensure that these effects were not the function of differential positive valence, I conducted an ANOVA on agreeableness ratings from the TIPI. No significant differences emerged, $F(2, 186) < 1$, *ns*. Mean ratings were 5.5 or higher in each condition, suggesting that across conditions the target was perceived as strongly agreeable.

I then tested whether the manipulation successfully induced the predicted differential humanization of the target. A one-way, three-level analysis of variance (ANOVA) showed a significant omnibus effect, $F(2, 186) = 10.16$, $p < .001$, $\eta^2_p = .10$ The ordering of conditions—from low-humanness to baseline to high-humanness—showed strong evidence for a linear trend, $F(1, 186) = 20.30$, $p < .001$. There was no evidence of a quadratic trend, $F < 1$, *ns*. Figure 4 graphically depicts the means. Humanization in the high-humanness condition was significantly greater than the baseline condition, $t(132) = 2.28$, $p < .05$, $d = .40$, while the baseline condition showed significantly greater humanization than the low-humanness condition $t(128) = 2.62$, $p = .01$, $d = .46$. Finally, humanization was significantly greater in the high-humanness than the low-humanness condition, $t(112) = 4.30$, $p < .001$, $d = .81$.

The central question of this study was whether increases in the humanization of a target cause increases in compassion toward the target. There was a significant omnibus effect of condition on compassion, $F(2, 186) = 2.96$, $p = .05$, $\eta^2_p = .03$. Although there was no evidence of a linear trend, $F < 1$, *ns*, instead a quadratic trend emerged, $F(1, 186) = 5.90$, $p < .05$. Figure 5 depicts the means. Consistent with prediction, the high-humanness condition evoked greater compassion toward the target than the baseline condition, $t(132) = 2.06$, $p < .05$, $d = .36$. Interestingly, the low-humanness condition, relative to the baseline condition, evoked increases in compassion as well, $t(128) = 2.06$, $p < .05$, $d = .36$. Finally, compassion evoked by the high-humanness and low-humanness condition were statistically indistinguishable, $t(112) = -.15$, *ns*.

This study shows that differential levels of humanness causally relate to compassion. In line with my prediction, high humanness, relative to a baseline, evoked greater levels of compassion. Interestingly, when the target was depicted as having low levels of humanness, compassion was also evoked, relative to the baseline. This latter effect may correspond to pity, conceptualized as compassion directed downward toward an inferior other (see Sayer, 2005). Here, the target is agreeable yet lacking uniquely human qualities, thus evoking concern and a desire to give aid and support, but to someone/something deemed helpless. This effect may explain certain paternalistic attitudes toward the mentally ill, who are often perceived as incompetent (Hinshaw & Stier, 2008; Teachman, Wilson, & Komarovskaya, 2006) and hence require control by an external authority rather than independence and rights (Ben-Zeev et al., 2012, Watson & Corrigan, 2001). This effect, and corresponding speculation, is worthy of further investigation. However, the results of the high-humanness condition (relative to the baseline) support my prediction that heightened levels of humanness evoke compassion, although there may be multiple routes to compassion: one from high-humanness, and another from low-humanness.

Study 5: Humanization, Compassion, Treatment-seeking, and Representations of the Other

Although increases in humanization evoke compassion, the underlying mechanism for this effect is unclear. In the next study, I directly test the notion that humanization of people with mental illness is associated with inclusion of people with mental illness into the representation of the self, and that this inclusive representation influences compassion. In addition, Study 5 tests the robustness of the link between humanization and treatment-seeking by using a different measure of this outcome. Specifically, I test whether humanization of people with mental illness predicts willingness to engage in a range of concrete treatment-seeking behaviors. Further, Study 5 seeks to more convincingly rule out a valence interpretation of the link between humanization and treatment-seeking by including a feeling thermometer (measuring warmth towards people with mental illness) as a covariate in the models. Finally, I test a more nuanced structural model, in which humanization activates a more inclusive self-representation (that includes those with mental illness), which influences compassion and eventuates in treatment-seeking tendencies.

Method

Participants

Participants were recruited for an online study through a national online retailer's data collection service and received monetary compensation in exchange for their participation. I specifically recruited participants who were neither formally diagnosed with a mental illness nor currently receiving treatment. Of the 234 participants that were recruited, 4 were excluded for not complying with the writing prompt (the "catch trial"). Of the remaining 230 participants (46.1 percent were female), 76.1 percent were European American, 7.0 Asian American, with 16.9 percent stated other ethnicities The average participant age was 33.10 years old (*SD* = 12.43). The mean political ideology was 3.11 (*SD* = 1.59) and the mean socioeconomic status was 2.62 (*SD* = .87).

Procedure

Participants were invited to participate in the study by clicking on a link that included a variety of measures. They were first asked to write about their opinions regarding mental illness and mental health. Thereafter, they completed a series of measures. After completing the study, participants were thanked, debriefed, and then awarded their compensation.

Measures

Humanization. Humanization was measured in the same way as in Study 4, but included humanity and animality items (Viki et al., 2006; Zebel et al., 2008). A composite measure was created with z-scores in which higher values indicate greater humanization, $\alpha = .91$.

Inclusion-of-outgroup-in-the-self. I measured the extent to which participants included members of the social category of mental illness into their self-representations by adapting a measure used in previous research that measures increases in overlap between the self and a group (Tropp & Wright, 2001). The current measure showed pairs of circles, with one circle representing the self, and the other circle representing an outgroup (here, specified to be people with mental illness). Each pair of circles showed varying degrees of overlap between the self and the other and was scored such that higher values indicated greater amounts of overlap between the self and people with mental illness.

Compassion. Compassion toward people with mental illness was measured with the same index employed in Study 4, $\alpha = .92$.

Feeling thermometer. I used a measure from the intergroup literature that assesses the degree of warmth felt toward members of another group (see Goren & Plaut, 2012). I adapted

this measure for warmth felt toward "people with mental illness."

Taking Action to Seek Care-Mental Health Version (TASC-MH). I developed a measure to assess people's willingness to engage in concrete treatment-seeking behaviors in the domain of mental healthcare. The four items employed were "*Do a google search about mental illness and treatment*," "*Visit a website about mental illness and treatment*," "*Attend a self-help group meeting about mental illness and treatment*," and "*Meet with a healthcare provider to talk about mental illness and treatment*." Participants rated their likelihood of engaging in these behaviors if they believed that they were having the symptoms of a mental illness. The anchors were 1 = "*not at all likely*" to 5 = "*extremely likely*."

A pilot sample rated these four items on level of commitment (1= *shows very little commitment*, 7 = *shows a great deal of commitment*), demonstrating that levels of commitment ascended monotonically (see Table 1). Moreover, paired-sample *t*-tests showed that each mean was significantly different from every other mean in the expected direction (all *p*s < .005). To reflect these ascending levels of commitment, as an item increased on this dimension, the specific item was multiplied by a positive integer (ascending from one to four). After these transformations, the four TASCI-MH items were summed. I standardized this measure to facilitate interpretation.[7]

Results and Discussion

To replicate the association between humanization and treatment-seeking (with a different measure), I used humanization to predict the TASC-MH. The bivariate regression was significant, $\beta = .29$, $p < .001$. Demonstrating that this effect was independent of valence, when the feeling thermometer measure was partialled, this relation remained significant, although it was attenuated in magnitude, $\beta = .19$, $p = .01$. Further testifying to the robustness of this relationship, the coefficient stayed significant when all demographic covariates were included in the regression model, along with the feeling thermometer, $\beta = .17$, $p < .05$

The next set of analyses tested whether inclusion-of-outgroup-in-the-self mediated the relation between humanization and compassion. Humanization predicted compassion ($\beta = .56$, $p < .001$). In light of the curvilinear results of Study 4, I tested for a quadratic effect of humanization on compassion, but this was not significant ($p = .10$), and thus I retained the linear form of this term. Completing the mediation analysis, humanization also predicted the mediator, inclusion-of-outgroup-in-the-self ($\beta = .43$, $p < .001$) and this mediator predicted compassion when controlling for humanization ($\beta = .24$, $p < .001$). The humanization term, though reduced in magnitude, was still significant ($\beta = .45$, $p < .001$), indicating partial mediation. The indirect effect was significant, as the bootstrapped confidence interval did not include zero (CI: .09, .35). The indirect effect remained significant when all demographic covariates were entered, including the feeling thermometer (CI: .01, .18). These results suggest that humanizing people with mental illness activates a more inclusive self-representation (which incorporates people with mental illness), and in turn increases compassion toward such individuals.[8]

The final set of analyses tested a comprehensive, and more nuanced model, in which humanization eventuates in treatment-seeking. Specifically, I tested the path whereby humanization activates inclusion-of-outgroup-in-the-self, which influences feelings of compassion, which in turn motivates perceivers' treatment-seeking (humanization → inclusion-of-outgroup-in-the-self → compassion → treatment-seeking). This model included a direct effect between humanization and compassion in light of partial mediation in the previous analysis. Figure 6 depicts this model and displays standardized path coefficients. Fit indices indicated that this model fit the data well: χ^2 (2) = 1.50, $p > .47$, *CFI* = 1.00, *RMSEA* = 0.00 (90 percent

confidence interval: 0.00 to .12), *SRMR* = .02. To gain confidence in this model, I also tested an alternative that reversed compassion and inclusion-of-outgroup in the self (humanization → compassion → inclusion-of-outgroup-in-the-self → treatment-seeking). Here, the inclusive mental representation, and not the emotion of compassion, is the proximate cause of treatment-seeking. The values of the fit-indices, however, undermine the viability of this model: χ^2 (2) = 35.40, $p < .001$, *CFI* = .82, *RMSEA* = 0.27 (90 percent confidence interval: 0.20 to .35), *SRMR* = .10. In light of these results, I retain my hypothesized psychological process model.

This study conceptually replicates the finding that humanization of mental illness eventuates in treatment-seeking. This relation emerged when using a different operationalization of treatment-seeking, a new measure (the TASC-MH) that includes a range of incremental concrete behaviors. In addition, these results further corroborate that the link between humanization and treatment-seeking cannot be explained away by positive valence. Importantly, this study also provides insight into how increases in humanization can result in increases in compassion. One intervening variable is a conception of the self that incorporates representations of the stigmatized other. Possessing this more inclusive self-representation leads to greater compassion toward the stigmatized other.

Finally, this study provides evidence for a more nuanced model in which humanizing people with mental illness eventuates in treatment-seeking. The humanization of people with mental illness, which leads to a more inclusive self-concept (that includes members of this stigmatized category), activates compassion toward these others, which in turn canalizes into perceivers' own willingness to seek treatment, should mental health problems befall the self. This suggests that the compassion toward stigmatized others influences a perceiver's own healthcare decisions, an effect emerging from an inclusive self-representation (that incorporates the stigmatized other) afforded by humanization.

Study 6: Examining the Causes of Humanizing People with Mental Illness, the Role of Behavior

The previous studies suggest that humanizing of people with mental illness has downstream effects on compassion and willingness to seek treatment. However, the reason that for this humanization is unclear. Study 3 manipulated humanization toward members of this category, but the manipulation contained several ingredients, any of which may have led to the humanization effect. The goal of this final study is to decompose Study 3's manipulation to address why humanization in this domain occurs. Study 6 uses the manipulation of Study 3, but includes two additional conditions. In addition to a no information baseline condition, as well as the successfully recovered condition (used in Study 3), Study 6 includes a condition in which the target with mental illness is described as having received no treatment as well as a condition is in which the target is described as receiving treatment that is inconsistently effective. Here, I test whether these different informational conditions elicit differential levels of humanization. I also test the possibility that treatment and recovery information do not drive the humanization effect, but that the principal determinant of this effect is the target's benign behavior. Further, I used an alternative operationalization of dehumanization (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996) and included a feeling thermometer to assess humanization independent of valence.

Method

Participants

Participants were recruited for an online study through a national online retailer's data collection service and received monetary compensation in exchange for their participation. I

specifically recruited participants who were neither formally diagnosed with mental illness nor receiving mental health treatment. Of the 243 participants that were recruited, 6 were excluded for substantial missing data. Of the remaining 237 participants (41.8 percent were female), 75.1 percent were European American, 5.9 percent were Asian American, with 19 percent stated other ethnicities The average participant age was 31.57 years old (*SD* = 11.49). The mean political ideology was 3.18 (*SD* = 1.64) and the mean socioeconomic status was 2.73 (*SD* = .85).

Procedure

Participants were invited to participate in the study by clicking on a link. Participants were randomly assigned to one of four conditions. In each condition they read a description of a person named Donald (employed in Study 3 and held constant across all four conditions in this study), from the point of view of his acquaintance. I included a no-further-information condition to serves as a baseline and three vignettes that were prefaced by information about Donald's mental health diagnostic and treatment status. The three types of information about diagnostic and treatment status were that the target was diagnosed with a mental illness but was *"not receiving treatment*," that he was diagnosed with a mental illness and had received treatment but *"the treatment has not always been effective*," and, finally, that he was diagnosed with a mental illness and *"has been successfully treated for as long as I can remember."*

After participants read the vignette, they were first asked to write a few sentences about their impressions of the target. This served as the catch trial. Thereafter, participants made ratings of the target and completed a demographic form. After exiting the study, participants were thanked, debriefed, and then they were awarded their compensation.

Measures

Humanization. Humanization was measured using the dehumanization subscale from the moral disengagement scale (Bandura et al, 1996). This four-item subscale includes such items as "*Some people deserve to be treated like animals*" and "*It is okay to treat badly somebody who behaved like a 'worm'*." Items are rated on five-point scale (1= *strongly disagree*, 5 = *strongly agree*). Each item was phrased in terms of Donald. To create an index such that higher values indicate increased humanization, thus making interpretation comparable to the previous studies, items were reverse-scored and then averaged, $\alpha = .86$.

Feeling thermometer. The feelings thermometer from Study 5 was used except the rated target was Donald.

Results and Discussion

To analyze a single humanization variable, but eliminate the potential confound between valence and humanization, I regressed humanization on the feeling thermometer and saved the residuals. These values represent that component of humanization that in independent of valence. I then conducted a one-way ANOVA with this variable as the dependent variable, treating each experimental condition as a level. This one-factor, four-level analysis led to a non-significant omnibus effect, $F(3, 233) = 1.87$, *ns*. However, I tested the hypothesis that irrespective of treatment information, the target with the mental illness label will be humanized relative to the baseline condition. To accomplish this, I coded the baseline condition as -3, and the other conditions as +1. This contrast was significant $t(233) = 2.27$, $p < .05$. To simplify interpretation and calculate Cohen's *d*, I collapsed the mental illness information conditions into a single category ($M = .08$, $SD = .92$) and compared this to the baseline condition ($M = -.26$, $SD = 1.18$), $F(1, 235) = 5.33$, $p < .05$, $d = .30$. Using this two condition coding to conduct an analysis of covariance (ANCOVA) on the original humanization measure, with the feeling thermometer as a covariate, yielded commensurate results.

These findings suggest that benign behavior of people labeled with mental illness, irrespective of treatment information, leads to humanizing. Importantly, this effect is independent of valence. Across conditions, this study presented participants with identical behavior: the common denominator of the three mental illness conditions was that the target was described as having a mental illness (although treatment information differed). Ultimately, it made no difference as to whether the target with mental illness was not receiving treatment, receiving treatment that worked inconsistently, or was receiving treatment that worked for a large span of time. These findings suggest that in the mental health domain, behavior matters most: as long as the perceiver is exposed to evidence of a target with mental illness behaving benignly, the target will be increasingly humanized. In other words, to shift humanizing (and likely other perceptions and attitudes) in the mental health arena, behavior has primacy. However, uncertainty remains about the enduring effects of interventions such as the one described in this study. It is possible that the memory traces of benign behavioral exemplars are unstable, and that stigma eventually reverts back to baseline levels. Future studies should include a post-manipulation follow-up to determine the longevity of informational manipulations like the one employed in this study.

General Discussion

This investigation illuminates humanization (and its antipode, dehumanization) by showing that this social cognitive process has important downstream implications for emotion and health behavior. I found that humanizing a stigmatized other (here focusing on the category of, and people with, mental illness) activates feelings of compassion toward such a target. Importantly, this investigation sheds light on why this occurs. Although theorists have posited the humanization leads to a more inclusive category that includes members of other social groups (Sayer, 2005), the current research provides empirical evidence that humanizing people from stigmatized outgroups influences the structure of the perceivers' self-concept. Humanizing the other leads to incorporating the other into the perceiver's self-representation. Importantly, humanization, via this inclusive representation, leads to particular feeling states that influence action tendencies. Specifically, humanization activates compassion, an emotion directed at providing support to and promoting the well-being of others (Goetz et al, 2010; Simon-Thomas et al., 2011). Interestingly, this investigation provides evidence that these compassionate feelings, which are typically directed outside the self, can re-direct to promote care-taking responses *toward the self*. In this way, humanizing a stigmatized health category can evoke motivations to care for oneself, should the self become threatened by illness. The intervening cognitive process influencing the pathway from humanization—that eventuates in self-care—is an inclusive self-representation, which evokes the emotion of compassion.

This psychological process model—in which humanization evokes an inclusive self/other representation, inducing compassion and, in turn, self-care tendencies—was mapped over the course of six studies. I found evidence for this model across a variety of operationalizations of the constructs (namely, humanization, compassion, and treatment-seeking), suggesting that these results are not specific to a particular type of measurement tool. Importantly, the effects were independent of measures of valence, a compelling alternative explanation of these findings. Further, I found evidence for the key outcome (treatment-seeking) when using a measure to assess seeing a professional (Mackenzie et al, 2004) or when using a measure that included incremental concrete behaviors that an individual could enact in the community (the Taking

Action to Seek Care or TASC measure). This research argues for a broader conception of (de)humanization in which (de)humanization is not only an intergroup (Leyens et al., 2007) and an interpersonal (Haslam et al., 2005) phenomenon but also has implications for individuals' health behavior.

(De)humanization and Health Behavior

A central motivation for this research was to discover if (de)humanization has implications for healthcare behavior. I found consistent evidence that humanizing the category and people with mental illness has downstream consequences for action tendencies in this domain. Whereas (de)humanization is typically linked with harmful consequences for stigmatized others (Haslam, 2006; Waytz & Epley, 2012), I investigated how humanizing stigmatized others may have positive, health-promoting outcomes.

Attributing humanness to the stigmatized may be an important determinant of healthcare decision-making for several reasons. In the case of mental illness stigma, the repugnance that perceivers experience toward this stigmatized category may be especially pronounced. Research suggests that, on average, the default response to encountering this category is dehumanization (Martinez et al., 2011) and persons with mental illness may be viewed as innately defective (see Martinez & Mendoza-Denton, 2011 for a review). Prior work shows that humanization activates approach tendencies (Martinez et al., 2011) as well as motivations to help others (Cuddy et al., 2007). To the extent that attributions toward people with mental illness can shift toward humanization, perceivers may be willing to seek help for themselves, should such care become warranted.

Interestingly, increasing humanization may require only that perceivers be exposed to a target with mental illness who behaves in benign ways. Study 6 suggests that perceivers are unaffected by the type of treatment (if any) that such a target receives, as long as behavior is benign. Given that the public is exposed to contradictory information about whether treatments actually work and is intermittently exposed to news stories about predatory pharmaceutical companies seeking to sell more medicine (see Hinshaw & Stier, 2008), people may be ambivalent, and even skeptical, about available mental health treatment options. However, behaviors are transparent, and exposure to the ordinary behavior of people who are typically feared may be enough to shift humanizing in a positive direction. Similar to research on members of racial minorities (Bodenhausen, Schwarz, Bless, & Waenke, 1995), repeated exposure to positive counterstereotypic *behavioral* exemplars may be all that is needed to impact stigma in this domain. The long-term stability of such shifted attitudes, however, needs to be further investigated.

To the extent that perceivers see those with mental illness as more human, either by exposure to new information or because of dispositional tendencies, perceivers may come to believe that others with mental illness are worthy of care. The reason for this is that humans qua humans are worthy of consideration, care, and concern (Sayer, 2005). Humanizing cognitions are important to the perceiver because perceivers may experience mental health concerns at a later time, as psychiatric epidemiology estimates that the lifetime prevalence rate of any mental illness is approximately 50 percent (see Hinshaw, 2007). Compassion directed toward the other, activated by humanization, may ultimately eventuate in care-taking responses directed toward the self, as the self now includes people with mental illness. In other words, by initially activating the perception that individuals with mental illness are humans too, should the perceiver experience mental health symptoms, he/she will be motivated to care for the self by seeking treatment.

Compassion as a Bridge Between Humanization and Health Behavior

These studies provide evidence that compassion is a mediator between humanization and health behavior. These findings cohere with theoretical approaches positing that emotions mediate cognition and action (Frijda et al., 1989; Weiner, quoted in Siegel & Shaughnessy, 1996). I find that humanization leads to an inclusive self-representation, that includes the stigmatized, and this mental representation activates the specific emotion of compassion. Whereas prior work has shown that compassion is associated with judgments that one is similar to seemingly unrelated others (Oveis et al., 2010), the current studies elucidate the nature of the underlying self-representation that allows this to happen. Further, the current studies enhance our understanding of the nature of compassion.

Compassion is likely a fluid emotion that can be transmitted to a variety of targets. Sayer (2005) theorizes that compassion is omni-directional: it can be directed laterally (toward an equal), upward (toward someone of higher status) or downward (toward someone of lower status). Compassion directed downward, toward lower-status others, may be pity, especially if the target is seen as agreeable (Fiske et al., 2006), and may elicit condescension. I found an effect in Study 4 in which the target was agreeable, and low levels of humanization evoked compassion toward the target. This result suggests that there are multiple pathways to compassion: one based on minimizing the humanness of an other and one based on ascribing humanness to an other. Should persons with mental illness be seen as warm, although incompetent and devoid of uniquely human qualities, pity and paternalism may be the dominant response (see Watson & Corrigan, 2001). However, increased humanizing may prompt a lateral compassion, promoting the belief that people with mental illness have autonomy and thus deserve rights and respect. This is an intriguing question that awaits future inquiry.

An important distinction in the current project is differentiating compassion directed toward the self, specifically if mental health concerns should arise, from the construct of self-compassion (Neff, 2003). Although the names are similar and may inspire some confusion, the underlying constructs differ. Neff's research on self-compassion (2003) describes a construct that involves being kind toward the self, seeing one's short-comings as shared by all people, and being mindful/non-judgmental toward the self in the face of one's mistakes. These attitudes are correlated with reduced anxiety and depression. Although this self-compassion is an important coping mechanism for dealing with stress and adversity, the current studies are investigating something else. Rather than examining a nonjudgmental acceptance of the self ("self-compassion"), the current studies investigate the feeling states that prompt seeking support and care-taking behaviors (Goetz et al, 2010). In the current research, the compassionate feelings that typically promote supportive action in the service of the well-being of an other (the traditional notion of compassion) are instead being directed toward the self, serving to promote the self's well-being through seeking care. This analysis suggests a modified notion of compassion, the implications of which should be studied further: compassion directed toward an other (*allo*-compassion) versus compassion directed toward the self (*auto*-compassion).

Limitations and Future Directions

The current studies take an important first step in mapping out what is going on in the mind of the perceiver but these studies do not actually measure the real-world behavior of seeking treatment in the community. The current research needs to be supplemented with field experiments. For example, after an initial laboratory session in which humanization and compassion are assessed, participants could be prospectively followed and clinic records could be audited (as investigators carefully protect disclosure of participants' health records). Such

studies would be resource-intensive, but feasible. Importantly, such studies would show how the psychological processes investigated in the current project relate to real-world health decisions.

Also, the current set of studies demonstrated effects for explicit humanization, but implicit humanization was not predictive of treatment-seeking. This too may be a consequence of a methodology that gives participants time to deliberate. In the context of an acute mental health crisis, characterized by time-pressure and extreme emotional arousal, implicit attitudes may exert effects. Although such a crisis cannot be easily created in the laboratory, the effects of implicit humanization could be investigated prospectively. Specifically, implicit attitudes could be assessed in a laboratory session and subsequent treatment-seeking could be assessed through administrative records.

Another area for future research involves employing biological measures to supplement the self-report measures that assess this investigation's key constructs: humanization and compassion. Although these are psychological phenomena, they also have neural and physiological concomitants. The current humanization measures could be supplemented with the measurement of individual differences in neural activation as participants reflect on a person with mental illness. It would be interesting to measure medial prefrontal cortical activation as an index of humanization (Harris, & Fiske, 2006), in response to a target labeled with mental illness, and use this variation to predict treatment-seeking behavior. Furthermore, compassion has reliable physiological correlates (Stellar et al., in press) and it would be interesting to measure heart rate deceleration as well as respiratory sinus arrhythmia (RSA) to index compassion toward people with mental illness.

Additional work should also determine whether the processes of humanization and compassion are unique to mental illness stigma or whether they hold more broadly. It has been shown, for example, that stigma adversely impacts medicine-adherence in HIV/AIDS (Rintamaki, Davis, Skripkauskas, Bennett, & Wolf, 2006). It is possible that the psychological processes involved with mental healthcare treatment-seeking—(de)humanization, self/other representations, and compassion—apply to a variety of stigmatized health domains. Other health conditions, especially those that perceivers associate with behavior (including obesity and diabetes), may also be influenced by the psychological processes examined in these studies. It would be informative to investigate mental illness stigma along side other stigmatized health conditions within the same study to determine the generality of these processes. Alternatively, mental illness stigma may be unique in terms of the underlying mental processes described in this paper. Future research should empirically determine the limits of the current psychological process model.

Whereas the current work focuses on the perceiver, future work should confront the other side of the stigma dynamic: the target. Although it is important to understand how people may make the initial decision to seek treatment, the psychology of people on the receiving end of stigma merits greater research attention. People who possess mental illness labels or access mental health treatment can internalize devaluation from perceivers, a process called self-stigma (Corrigan, 2004). Research should determine how dehumanization affects targets with a mental illness label. Basic social psychological research suggests that manipulating ostracism and rejection can make a target feel less human (Bastian & Haslam, 2010). Might the stigma, devaluation and rejection experienced by targets with mental illnesses make them feel subhuman? Based on the current research, feeling dehumanized may disrupt adherence to health regimens, as I find a consistent humanization/treatment-seeking linkage. Future research should rigorously examine this issue. Further, the shame associated with having a mental illness

(Hinshaw, 2007) may further lead to a target's sense of feeling less than others, if not subhuman. Moreover, efforts to conceal one's stigmatized identity (Quinn & Chaudoir, 2009), intensified by felt shame and subhuman status, may also exacerbate treatment avoidance and non-adherence. The present research, which has shed light on some of the constructs that are important for the perceiver, may guide and inform research that examines the target's perspective, a worthwhile area for future inquiry.

Conclusion

Although many effective treatments are available for mental health conditions, services remain underutilized. This treatment gap may in part be driven by the stigma associated with accessing mental healthcare (Ben-Porath, 2002; Corrigan, 2004). The current investigation provides an explanation of this state of affairs. Specifically, this research highlights the importance of humanizing people with mental illness, a process that creates a self-representation that incorporates these stigmatized targets, and thereby evokes compassion toward them. These feelings of compassion toward stigmatized others can redirect toward the self, increasing perceivers' own willingness to seek treatment should he/she experience symptoms. In this way, humanizing targets with mental illness may not only be a potent force for changing stigma, but may also influence underlying self-representations and emotions, thereby motivating people to pursue treatment when it is needed.

References

Akerlof, G.A., & Kranton, R.E. (2010). *Identity economics: How our identities shape our work, wages, and well-being.* Princeton: Princeton University Press.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Bandura, A., Barbaranelli, C., Caprara, G., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71,* 364–374.

Bastian, B., & Haslam, N. (2010). Excluded from humanity: The dehumanising effects of social ostracism. *Journal of Experimental Social Psychology, 46,* 107–113.

Bastian, B., Loughnan, S., Haslam, N., & Radke, H. (2012). Don't mind meat? The denial of mind to animals used for human consumption. *Personality and Social Psychology Bulletin, 38,* 247–256.

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., et al. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, *72*, 105–118.

Ben-Porath, D. D. (2002). Stigmatization of individuals who receive psychotherapy: An interaction between help-seeking behavior and the presence of depression. *Journal of Social & Clinical Psychology, 21,* 400-413.

Ben-Zeev, D., Corrigan, P.W., Britt, T.W., & Langford, L. (2012). Stigma of mental illness and service use in the military. *Journal of Mental Health, 21,* 264-273.

Blanco, C., Okuda, M., Wright, C., Hasin, D. S., Grant, B. F., Liu, S. M., et al. (2008). Mental health of college students and their non-college-attending peers: Results from the National Epidemiologic Study on Alcohol and Related Conditions. *Archives of General Psychiatry, 65,* 1429–1437.

Bodenhausen, G. V., Schwarz, N., Bless, H., & Waenke, M. (1995). Effects of atypical exemplars on racial beliefs: Enlightened racism or generalized appraisal*? Journal of Experimental Social Psychology, 31,* 48-63.

Carman, M. M. (2009). *The compassionate imagination in "Paper Towns" and "King Dork"* (Master's thesis). Retrieved from Dissertations and Theses database. (UMI No. 1470886)

Corrigan, P. (2004). How stigma interferes with mental health care. *American Psychologist, 59,* 614–625.

Costello, K., & Hodson, G. (2010). Exploring the roots of dehumanization: The role of animal human similarity in promoting immigrant humanization. *Group Processes and Intergroup Relations*, *13*, 3–22.

Côté, S., Gyurak, A., & Levenson, R.W. (2010). The ability to regulate emotion is associated with greater well-being, income, and socioeconomic status. *Emotion, 10*, 923-933.

Crocker, J., Major, B., & Steele, C. (1997). Social stigma. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 4). Boston: McGraw–Hill.

Cuddy, A. J. C., Rock, M., & Norton, M. I. (2007). Aid in the aftermath of Hurricane Katrina: Inferences of secondary emotions and intergroup helping. *Group Processes and Intergroup Relations, 10,* 107-118.

Demoulin, S., Rodríguez-Torres, R., Rodríguez- Pérez, A., Vaes, J., Paladino, M. P., Gaunt, R., et al. (2004) Emotional prejudice can lead to infrahumanization. *European Review of Social Psychology*, *15*, 259–296.

Epley, N., Waytz, A., & Cacioppo, J.T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review,114,* 864–886.

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J.T. (2008). Creating social connection through inferential reproduction: Loneliness and perceived agency in gadgets, gods, and greyhounds. *Psychological Science, 19,* 114–120.

Ekman, P. (2010). Darwin's compassionate view of human nature. *Journal of the American Medical Association, 303,* 557-558.

Fiske, S. T., Cuddy, A., & Glick, P. (2006). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11,* 77–83.

Frijda, N. H., Kuipers, P., & ter Schure, *E.* (1989). Relations among emotion appraisal and emotional action readiness. *Journal of Personality and Social Psychology, 57,21*2-228.

Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin, 136,* 351–374.

Goren, M.J., & Plaut, V.C. (2012). Identity form matters: White racial identity and attitudes toward diversity. *Self and Identity, 11,* 237–254.

Gosling, S. D., Rentfrow, P. J., & Swan, W. B., Jr. (2003). A very brief measure of the big five personality domains. *Journal of Research in Personality, 37,* 504-528.

Gray, H. M., Gray, K., & Wegner, D. M. (2007) Dimensions of mind perception. *Science, 315,* 619.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197-216.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Harris, L.T., & Fiske, S.T. (2006). Dehumanizing the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science, 17,* 847–853.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*, 252–264.

Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you: Attributing humanness to self and others. *Journal of Personality and Social Psychology, 89,* 937–950.

Haslam, N., Loughnan, S., Reynolds, C., & Wilson, S. (2007). Dehumanization: A new perspective. *Social and Personality Psychology Compass*, *1*, 409-422.

Haslam, N., Loughnan, S., Kashima, Y., & Bain, P. (2008). Attributing and denying humanness to others. *European Review of Social Psychology*, *19*, 55–85.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*, 408-420.

Hinshaw, S.P. (2007). *The mark of shame: Stigma of mental illness and an agenda for change*. New York: Oxford University Press.

Hinshaw, S. P., & Stier, A. (2008). Stigma as related to mental disorders. *Annual Review of Clinical Psychology, 4,* 367-393.

Hodson, G., & Costello, K. (2007). Interpersonal disgust, ideological orientations, and dehumanization as predictors of intergroup attitudes. *Psychological Science, 18,* 691–698.

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology, 97,* 963-976.

Keltner, D. (2009). *Born to be good: The science of a meaningful life.* New York: W.W. Norton.

Keltner, D., Marsh, J., & Smith, J. A. (Eds.). (2010). *The compassionate instinct: The science of human goodness.* New York: W. W. Norton & Company.

Kessler, R.C., Berglund, P.A., Demler, O., Jin, R., & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry, 62,* 593-602.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.

Kohn, R., Saxena, S., Levav, I., Saraceno, B. (2004). The treatment gap in mental health care. *Bulletin of the World Health Organization, 82,* 858–866.

Kozak, M. J., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology, 90*, 543–555.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of board of scientific affairs' advisory group on the conduct of research on the internet. *American Psychologist, 59*, 105-117.

Kwan, V. S. Y., & Fiske, S. T. (2008). Missing links in social cognition: The continuum from nonhuman agents to dehumanized humans. *Social Cognition, 26,* 125–128.

Lane, K.A., Banaji, M.R., Nosek, B.A., & Greenwald, A.G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) about the method. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Progress and controversies* (pp. 59–102). New York: Guilford Press.

Leyens, J.-P., Demoulin, S., Vaes, J., Gaunt, R., & Paladino, P. M. (2007). Infrahumanization: The wall of group differences. *Social Issues and Policy Review, 1,* 139–172.

Link, B. G., & Phelan, J. C. (2006). Stigma and its public health implications. *Lancet, 367,* 528–529.

Link, B. G., Cullen, F. T., Frank, J., & Wozniak, J. F. (1987). The social rejection of former mental patients: Understanding why labels matter. *American Journal of Sociology, 92,* 1461–1500.

Link, B. G., Phelan, J. C., Bresnahan, M., Stueve, A., & Pescosolido, B. A. (1999). Public conceptions of mental illness: Labels, causes, dangerousness, and social distance. *American Journal of Public Health, 89,* 1328-1333.

Loughnan, S., & Haslam, N. (2007). Animals and androids: Implicit associations between social categories and nonhumans. *Psychological Science, 18*, 116–121.

Loya, F., Reddy, R., & Hinshaw, S.P. (2010). Mental illness stigma as a mediator of differences in Caucasian and South Asian college students' attitudes toward psychological counseling. *Journal of Counseling Psychology, 57*, 484-490.

Mackenzie, C. S., Knox, V. J., Gekoski, W. L., & Macaulay, H. L. (2004). An adaptation and extension of the Attitudes toward Seeking Professional Psychological Help scale. *Journal of Applied Social Psychology, 34,* 2410–2435.

Marcotte, D.E., & Markowitz, S. (2009). A cure for crime? Psycho-pharmaceuticals and crime trends. *National Bureau of Economic Research, working paper number 15354.*

Martinez, A.G., & Mendoza-Denton, R. (2011). The prospect of plasticity: Malleability views of group differences and their implications for intellectual achievement, mental/behavioral health, and public policy. *Social Issues and Policy Review, 5,* 137-159.

Martinez, A. G., Piff, P. K., Mendoza-Denton, R., & Hinshaw, S. P. (2011). The power of a label: Mental illness diagnoses, ascribed humanity, and social rejection. *Journal of Social & Clinical Psychology, 30*, 1-23.

Mehta, S., & Farina, A. (1997). Is being "sick" really better? Effect of the disease view of mental disorder on stigma. *Journal of Social and Clinical Psychology, 16,* 405–419.

Mendoza-Denton, R. (2008). Stigma. In W.A. Darity (Ed.), *International Encyclopedia of the Social Sciences, 2nd edition (Vol., 8, pp. 146-148).* Detroit: Macmillan Reference USA.

Neff, K. D. (2003). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity, 2,* 85–102.

Oveis, C., Horberg, E. J., & Keltner, D. (2010). Compassion, pride, and social intuitions of self–other similarity. *Journal of Personality and Social Psychology, 98,* 618–630.

Paladino, M.-P., Leyens, J.-P., Rodriguez-Torres, R., Rodriguez-Perez, A., Gaunt, R., & Demoulin, S. (2002). Differential association of uniquely and non-uniquely human emotions to the ingroup and the outgroup. *Group Processes and Intergroup Relations, 5,* 105–117.

Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5,* 411–419.

Post, R. M. (1992). Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *American Journal of Psychiatry, 149,* 999–1010.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36,* 717-731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879-871.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. E (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67,* 741-763.

Quinn, D.M., & Chaudoir, S.R. (2009). Living with a concealable stigmatized identity: The impact of anticipated stigma, centrality, salience, and cultural stigma on psychological distress and health. *Journal of Personality and Social Psychology, 97,* 634-651.

Rintamaki, L. S., Davis, T. C., Skripkauskas, S., Bennett, C. L., & Wolf, M. S. (2006). Social stigma concerns and HIV medication adherence. *AIDS Patient Care & STDS, 20,* 359-368.

Sayer, A. (2005). *The Moral Significance of Class*. Cambridge: Cambridge University Press.

Sieff, E. M. (2003). Media frames of mental illnesses: The potential impact of negative frames. *Journal of Mental Health, 12,* 259–269.

Siegel, J., & Shaughnessy, M.F. (1996) An interview with Bernard Weiner. *Educational Psychology Review, 8,* 165-174.

Simon-Thomas, E.R., Godzik, J., Castle, E., Antonenko, O., Ponz, A., Kogan, A., & Keltner, D. (in press). An fMRI study of caring *vs* self-focus during induced compassion and pride. *Social Cognitive and Affective Neuroscience.*

Smith, A. (2002). *The theory of moral sentiments.* New York: Cambridge University Press. (Original work published 1759)

Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37,* 1660-1672.

Stellar, J. E., Manzo, V. M., Kraus, M. W., & Keltner, D. (in press). Compassion and class: Socioeconomic factors predict compassionate responding, *Emotion*.

Szasz, T. (2007). *The medicalization of everyday life.* New York: Syracuse University Press.

Teachman, B., Wilson, J., & Komarovskaya, I. (2006). Implicit and explicit stigma of mental illness in diagnosed and healthy samples. *Journal of Social and Clinical Psychology, 25,* 75 – 95.

Tropp, L. R., & Wright, S. C. (2001). Ingroup identification as the inclusion of ingroup in the self. *Personality and Social Psychology Bulletin, 27,* 585–600.

Viki, G. T., Winchester, L., Titshall, L., Chisango, T., Pina, A., & Russell, R. (2006). Beyond secondary emotions: The infrahumanization of outgroups using human-related and animal-related words. *Social Cognition, 24,* 753-775.

Wahl, O. F. (1995). *Media madness: Public images of mental illness.* New Brunswick, NJ: Rutgers University Press.

Wang, P. S., Simon, G., & Kessler, R. C. (2003). The economic burden of depression and the cost-effectiveness of treatment. *International Journal of Methods in Psychiatric Research, 12,* 22–33.

Watson, A.C., & Corrigan, P.W. (2001). *The impact of stigma on service access and participation: A guideline developed for the behavioral health recovery management project.* [Brochure].

Waytz, A., & Epley, N. (2012). Social connection enables dehumanization. *Journal of Experimental Social Psychology, 48,* 70–76.

Waytz, A., Epley, N., & Cacioppo, J.T. (2010). Social cognition unbound: Psychological insights into anthropomorphism and dehumanization. *Current Directions in Psychological Science, 19,* 58–62.

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Science, 14*, 383-388.

Waytz, A., & Mitchell, J. P. (2011). Two mechanisms for simulating other minds: Dissociations between mirroring and self-projection. *Current Directions in Psychological Science, 20,* 197-200.

Zebel, S., Zimmermann, A., Viki, G. T., & Doosje, B. (2008). Dehumanization and guilt as distinct but related predictors of support for reparation policies. *Political Psychology, 29*, 193-219.

Footnotes

1. This formulation about androids raises questions about the universality and cultural-specificity of (de)humanization. It may be that a certain degree of technological development is required for mechanistic dehumanization (that is, likening individuals to androids). This would imply that only cultures that employ machines could dehumanize in this way. Further, if (de)humanization is culturally contingent, this implies that as societies develop, new modes of (de)humanization will also develop. Fully understanding the denial and attribution of humanity may necessitate greater reliance on qualitative historical and ethnographic data.

2. I also included a parallel implicit measure of humanization by employing an implicit association test or IAT (Greenwald, McGhee, & Schwartz, 1998). Here the category/target stimuli were mental illness labels (versus physical illness labels) and the attribute stimuli were human (versus animal) words. Attribute stimuli were the identical to the current explicit humanization measure and IAT scoring employed the revised algorithm (Greenwald, Nosek, & Banaji, 2003). Implicit (de)humanization was not related to treatment-seeking. Moreover, implicit (de)humanization did not moderate the effect of explicit humanization on willingness to seek treatment. These null results accord with current theorizing that posits that implicit attitudes better predict "heat of the moment" behaviors (Lane, Banaji, Nosek, & Greenwald, 2007) rather than more deliberative decisions. To the extent that treatment-seeking is deliberative (or contains elements of deliberation), this principle likely holds.

3. All bootstrapped mediation models employ bias corrected and accelerated confidence intervals with 10,000 resamples.

4. An important issue is the sense in which the target was humanized. I maintain that the target was construed as human in the sense of part of an inclusive humanity, and not weak or *"only human.*" I also collected data using an 8-item index of how weak the target was perceived to be ($\alpha = .84$). An "only human" interpretation predicts that the target with mental illness would be seen as weaker than the comparison condition. However, the opposite was observed, $F(1,82) = 5.27$, $p < .05$, $d = .51$.

5. I also tested a potential moderator of these effects: anti-egalitarianism or social dominance orientation (SDO). I employed the standard (explicit) SDO measure (Pratto, Sidanius, Stallworth, & Malle, 1994, Appendix C) as well as an IAT. The implicit measure's target categories were *hierarchy* versus *equality* and the attribute category was valence (here I employed a subset of positive and negative exemplars from Greenwald et al., 1998). Explicit SDO did not moderate the effect of condition on humanization. Moreover, the implicit SDO measure did not moderate the effect of condition on humanization. Finally, the three-way interaction (condition X explicit SDO X implicit SDO) was not significant. A similar pattern of null results was obtained when I repeated these analyses on the outcome variable of treatment-seeking. Stated simply, these analyses yielded null results straight down the line (all *p*s > .30). In sum, SDO—measured explicitly or implicitly—did not moderate any effects.

6. To rule out competing models, two alternative structural models were tested, both of which indicated poor fit (Figure 3 illustrates these models diagrammatically). The first model shows the manipulation evoking compassion, which drives humanization, which in turn motivates treatment-seeking. Ruling out this model, indices indicated poor fit: χ^2 (2) = 8.08, $p < .05$, *CFI* = .70, *RMSEA* = 0.19 (90 percent confidence interval: 0.07 to .34, *SRMR* = .08. The second model shows the manipulation activating humanization and compassion simultaneously, which in turn lead to treatment-seeking. Ruling out this model, indices indicated poor fit on the whole: χ^2 (1) = 3.82, $p = .05$, *CFI* = .86, *RMSEA* = 0.18 (90 percent confidence interval: 0.00 to .39), *SRMR* = .06. In light of these results, I retain the hypothesized psychological process model.

7. I also repeated the analyses by summing the TASC-MH items (rather than weighting and then summing the items). The pattern and significance of the results remained the same.

8. A mental representation that includes the outgroup in the self is not the only possible social cognitive mediator of the humanization-compassion link. In light of this, I tested another possible self/outgroup representation: separate-subgroups/superordinate-ingroup. This was operationalized as a pictorial representation that showed two separate, non-overlapping circles, one for the self and another for the outgroup (in this case people with mental illness). However, these two separate circles had a larger circle surrounding them, to depict a common or shared superordinate category. I found that, with covariates in the model, this representation was negatively related to humanization ($\beta = -.17$, $p < .05$) but did not mediate the link between humanization and compassion.

Table 1

Rated Level of Commitment for Taking Action to Seek Care-Mental Health Version (TASC-MH), Pilot Study

	Google Search	Website	Self-help Group	Meet with Healthcare Provider
Means:	3.56	4.07	5.87	6.41
(*sd*s)	(1.72)	(1.66)	(1.33)	(1.09)

Figure Captions

Figure 1: Mediational model from Study 2. Individual differences in humanizing the category of mental illness predict increases in compassion toward targets with mental illness, which in turn predicts increases in perceivers' treatment-seeking tendencies. Standardized regression coefficients are displayed. *Note: *** indicates $p < .001$ and ** indicates $p < .01$.*

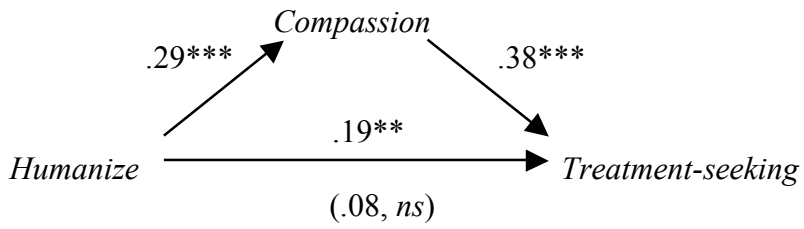
Figure 2: Path model tested in Study 3. Experimentally-induced humanization of a recovered target with mental illness evokes compassion toward the target, which in turn increases perceivers' treatment-seeking. Standardized path coefficients are displayed. *Note: * indicates $p < .05$ and ** indicates $p < .01$.*

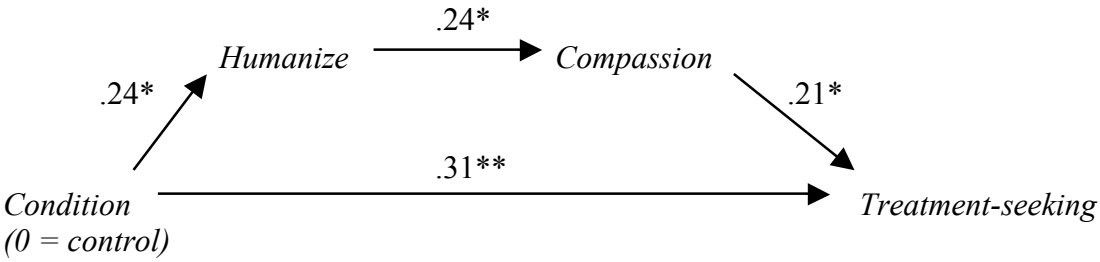
Figure 3: Two alternative structural models tested in Study 3.

Figure 4: Humanization of target as a function of experimental condition in Study 4. Error bars show standard errors.

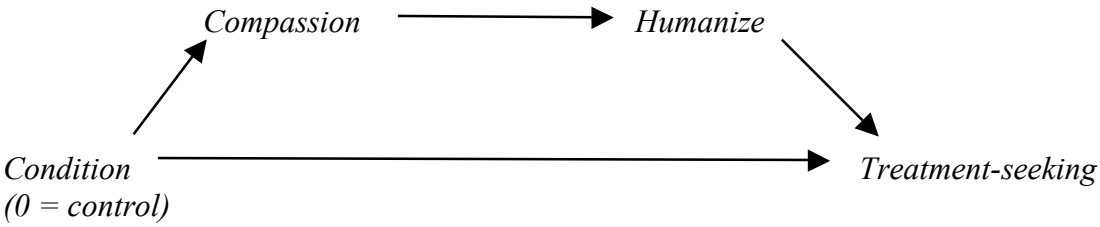
Figure 5: Compassion toward target as a function of experimental condition in Study 4. Error bars show standard errors.

Figure 6: Path Model from Study 5 in which humanization of a targets with mental illness influences inclusion-of-outgroup-in-the-self, which in turn influences compassion toward people with mental illness, thereby increasing perceivers' treatment-seeking (measured by the TASC-MH). All path coefficients are significant at $p < .001$.

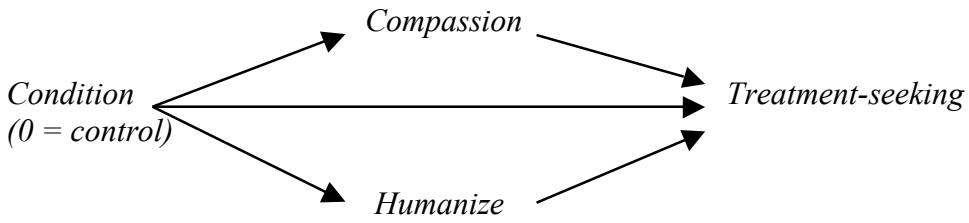

Compassion
.29***
.38***
.19**
Humanize
Treatment-seeking
(.08, ns)


.24*
Humanize
Compassion
.24*
.21*
.31**
Condition
(0 = control)
Treatment-seeking

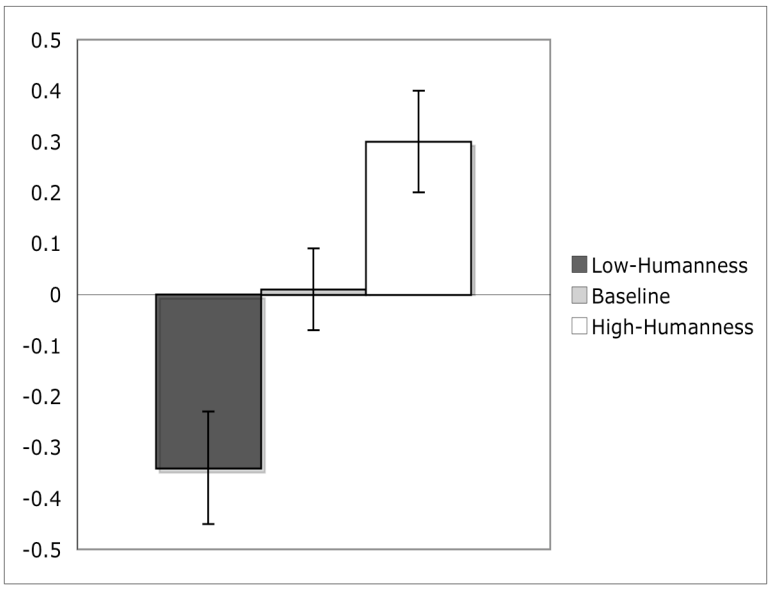
Study 3: Alternative model 1



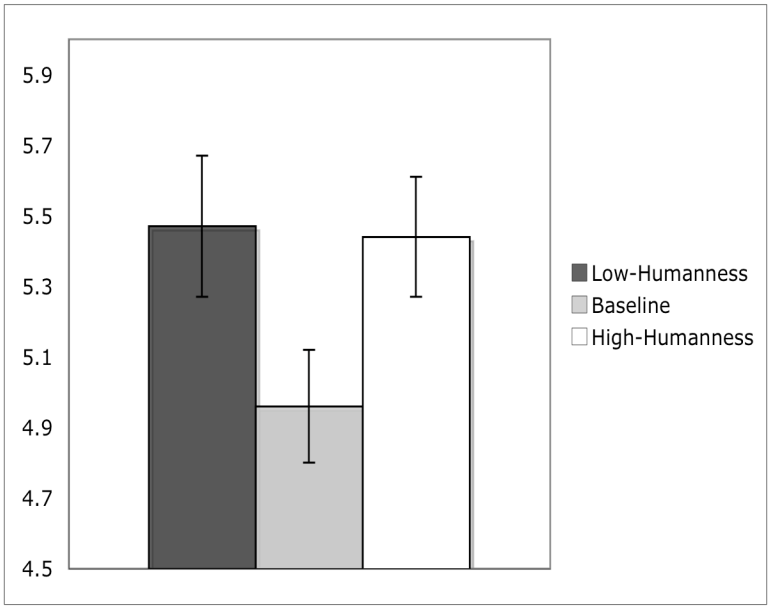
Study 3: Alternative model 2

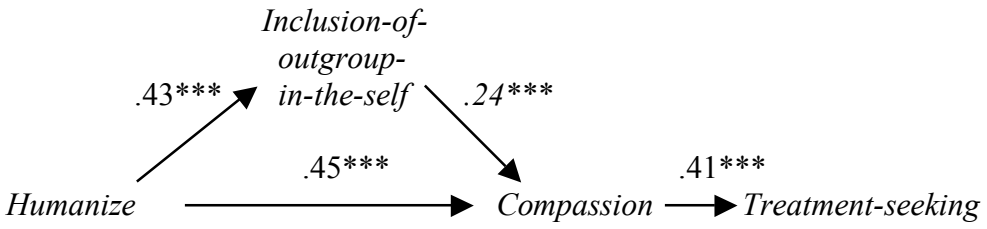


Humanization of Target as a Function of Condition



Compassion Toward Target as a Function of Condition




Inclusion-of-outgroup-in-the-self
.43***
.24***
.45***
.41***
Humanize
Compassion
Treatment-seeking