Reframing mental illness: The role of essentialism on perceived treatment efficacy and stigmatization

Danielle Peters[1] · David Menendez[1] · Karl Rosengren[2]

Abstract

People believe that treatments for illnesses are effective when they target the cause of the illness. Prior work suggests that biological essentialist explanations of mental illness lead people to prefer medications or other pharmacological treatments. However, prior work has not distinguished between biological and essentialist explanations. In three studies (total n = 517), we presented adults with vignettes about an individual with an artificial mental illness and manipulated the descriptions to emphasize or de-emphasize essentialist characteristics. Critically, none of the vignettes made reference to a biological basis for the disorder. Participants rated their willingness to interact with the person described in the vignettes and how effective they believed drug treatment and talk therapy would be on the mental illness. Across the three studies, describing mental illness with an essentialist framing led participants to think drug treatments would be more effective, but there was no effect for stigma or perceived effectiveness of talk therapy. This effect appears to be mediated by how much participants essentialized individuals with the disorder. The first framing that participants encountered seemed to shape their reasoning for the remainder of the study, even if they saw conflicting framing later on. The framing manipulation had similar effects for individuals with and without a mental illness. Results suggest that it is important to consider how mental illness is framed to the general public as it might impact people's treatment preferences.

Keywords Essentialism · Social cognition · Perceive Treatment Effectiveness · Stigma · mental illness

One in five Americans suffer from a mental illness (National Institute of Mental Health, 2017), but less than half of the people who had experienced a mental illness in the past year received treatment (National Institute of Mental Health, 2017). Patient preferences play a significant role in treatment. Past research has shown that patients with depression who receive treatment that aligns with their own preference for medication or psychotherapy are more likely to initiate and adhere to treatment (Raue, Schulberg, Heo, Klimstra, & Bruce, 2009) and have higher remission rates and lower levels of depression (Kocsis et al., 2009). Given this association between preference and adherence to treatment, it is critical to understand why people might prefer certain treatments or believe them to be more effective. One possible factor that could influence beliefs and attitudes about treatments might be essentialist beliefs.

Essentialism and mental illness

Psychological essentialism refers to the notion that people believe that categories have an underlying property (an essence) that determines category membership (Medin & Ortony, 1989). People tend to hold essentialist beliefs about simple categories such as "dogs" or "vegetables," but also about complex social categories such as race, gender, and sexual orientation (Dar-Nimrod & Heine, 2011). Essentialist beliefs encompass various subcomponents. People who hold essentialist beliefs about a category are likely to believe that knowing that someone is a member of a certain category is greatly informative (informativeness), that the category has existed across history with very few changes (historical invariance), that the category has all-or-none boundaries (discreteness), and that members of the category are highly similar to one another (uniformity; Gelman, 2003, 2004; Haslam & Ernst, 2002). They are also likely to believe that it is difficult or impossible for a category member to lose membership (immutability), that there are certain characteristics necessary to

✉ David Menendez
  dmenendez@wisc.edu

[1] Department of Psychology, University of Wisconsin Madison, 1202 West Johnson Street, Madison 53706, WI, USA

[2] Department of Psychology and Department of Brain and Cognitive Sciences, University of Rochester, 500 Wilson Blvd, Rochester 14611, NY, USA

be a category member (necessary features), that category membership is due to an inherent underlying reality (inherence), and that the category is naturally occurring rather than socially constructed (naturalness; Gelman, 2003; Haslam & Ernst, 2002).

Many people hold essentialist beliefs about mental illness, believing that the disorder is caused by something inside the person that is core to their identity (Ahn, Flanagan, Marsh, & Sanislow, 2006). This might be problematic as essentialist beliefs of mental illness affect both stigmatization and views on treatment (Dar-Nimrod & Heine, 2011). One reason why people might hold essentialist beliefs about mental illness is that explanations of mental illness frequently attribute internal biological factors as the causes. This is not unique to mental illness, as in other domains people often appeal to an underlying biological cause such as the heart, other organs, DNA, or blood to explain category membership (Balkcom, Alogna, Curtin, Halberstadt, & Bering, 2019; Roberts & Gelman, 2015; Waxman, Medin, & Ross, 2007). Biological framing of mental illness may also serve to activate essentialist reasoning and lead to negative outcomes such as stigma (Loughman & Haslam, 2018).

Biology, essentialism, and perceived treatment efficacy

Biological explanations of mental illness influence people's beliefs about the efficacy of treatment in general. Lebowitz, Rosenthal, and Ahn, (2016) found that, when reading vignettes about children with ADHD, reading a biological explanation of the disorder decreased stigma but increased doubt about treatment efficacy. Additionally, Marsh and Romano (2016) found that people often think that drug treatment would be more effective for symptoms perceived as medically based and that talk therapy would be more effective for symptoms perceived as psychological in nature. Beliefs about the etiology of mental illness influence beliefs about what kind of treatment is appropriate. That is, if the person believes that symptoms are due to an internal, biological cause, they might perceive treatments that modify something inside the person to be more effective. Yopchick and Kim (2009) found that when making judgements about treatment efficacy, people consider the root cause of the mental illness to be most important. If the root cause of the illness was described as biological, people believed that drug therapy would be more effective, and if the root cause was described as psychological, they believed psychotherapy would be more effective at treating the illness. Similarly, Lebowitz and Appelbaum (2017) found that reading genetic explanations of addiction increased confidence in pharmacotherapy and decreased confidence in psychotherapy. In addition, Phelan, Yang, and Cruz-Rojas (2006) found that belief in a biological cause of mental illness was related to greater endorsement of hospitalization and medication, but lower expectations that a mental health professional could help treat the illness.

Although biological explanations of mental illness frequently attribute internal factors as the cause of the illness, this need not be the case, as individuals can hold essentialist beliefs about nonbiological categories (e.g., art; Gelman & Bloom, 2000). Therefore, it could be that appealing to internal causes, without mentioning a biological cause, might promote an essentialist view of mental illness. Appealing to an internal cause might also influence treatment decisions (Kim & LoSavio, 2009). For example, Schroder, Dawood, Yalch, Donnellan, and Moser (2015) found that people who had an essentialist view were more likely to prefer medication alone over psychotherapy or a combined treatment. If this is true, then describing mental illnesses in essentialist terms might lead people to prefer treatments (such as medication) that have effects internally, even if a biological cause is not mentioned in the explanation.

Biology, essentialism, and stigma

In addition to influencing beliefs about treatment, essentialist beliefs about mental illness might also play a role in stigmatization (Dar-Nimrod & Heine, 2011). Stigmatization is the act of distinguishing and labeling differences between humans, normally with a negative connotation, and separating "us" and "them" based on those social differences (Link & Phelan, 2001). The consequences of stigmatization are wide ranging, including loss of self-esteem, job discrimination, and avoidance of treatment (Rüsch, Angermeyer, & Corrigan, 2005).

Essentialist views about social categories such as race, gender, and sexual orientation have been shown to be related to prejudice and stereotyping (Dar-Nimrod & Heine, 2011). People who hold essentialist beliefs are more likely to support legislation that enhances boundaries between social groups, and these beliefs can be manipulated by providing information that either confirms or disconfirms the essentialist belief (Roberts, Ho, Rhodes, & Gelman, 2017). Researchers have also found that people who hold essentialist beliefs about mental illness have more stigmatizing attitudes about people with a mental illness (Howell, Weikum, & Dyck, 2011). However, some studies have failed to find an association between different subcomponents of essentialism and stigmatization of individuals with mental illness (Marsh & Shanks, 2014).

Biological explanations of mental illness and essentialist beliefs may interact in complex ways to influence stigmatization. Biological explanations of mental illness may reduce stigma by shifting the blame from the individual to biological factors outside of the individual's control (Lebowitz et al., 2016). Indeed, Goldstein and Rosselli (2003) found that people who believe that depression is caused by biological factors were less likely to blame people with depression for their illness. However, biological explanations may increase the stigma toward people with a mental illness (Phelan, 2002). For example, Walker and Read (2002) found that people who heard a biological explanation of

schizophrenia believed that people with schizophrenia were more dangerous and unpredictable than those who heard a psychosocial (nonessentialist) explanation. Even in the same studies there have been conflicting findings. For example, Breheny (2007) found that providing a genetic explanation for schizophrenia decreased stigmatization, but that providing a genetic explanation for depression increased stigmatization. These results suggest that the relation between stigma and biological or essentialist explanations is complex, as biological essentialist explanations could decrease some components of stigma (such as blame), but increase others (such as social distancing; Haslam & Kvaale, 2015). Currently, it is difficult to predict when one of the outcomes will occur. One issue with the prior research is that many of these studies combine essentialist and biological information. This combination makes it difficult to pinpoint whether different types of information have either beneficial or negative effects. In our study, we hope to shed light on this relation by examining the effects of essentialist information on its own.

Essentialism among people with a mental illness

Most of the literature on essentialism and mental illness has focused on the beliefs of the general public, but less attention has been given to how individuals with a mental illness respond to essentialist explanations. People who have received psychiatric services have more accepting attitudes toward others with a mental illness (Segal, Kotler, & Holschuh, 1991; Walker & Read 2002), and so might interpret essentialist information differently. It has been found that people who belong to a stigmatized group often respond differently to essentialist information. For example, although essentialist views of sexual orientation (such as "born this way") relate to lower stigma among heterosexual individuals (Haslam & Levy, 2006; Haslam, Rothschild, & Ernst, 2002), essentialist views of sexual orientation have mixed effects among homosexual and bisexual individuals (Morandini, Blaszczynski, Costa, Godwin, & Dar-Nimrod, 2017; Morandini, Blaszczynski, Ross, Costa, & Dar-Nimrod, 2015; Morton & Postmes, 2009).

One important context in which people with a mental illness may encounter essentialist information is the context of treatment. Biological explanations of psychological symptoms, when viewed through an essentialist lens, may influence people's beliefs about the course and treatment of their illness. In one study that examined the effects of biological explanations among people with mental illness, Kemp, Lickel, and Deacon (2014) randomly assigned individuals who have had a depressive episode to either being told that the episode was due to a neurochemical imbalance (i.e., a biological explanation) or not. Participants who were told that their depression was caused by a neurochemical imbalance displayed increased perceived stigma, greater pessimism about their diagnosis and treatment, and

lower perceived ability to regulate their own negative mood states. These results suggest that biological explanations for mental illness may negatively affect those who suffer from mental illnesses.

Similarly, past research has also demonstrated that people with depression who endorse biological explanations for their symptoms display greater prognostic pessimism, the belief that mental illnesses are likely to be stable over time and difficult or impossible to treat (Lebowitz, Ahn, & Nolen-Hoeksema, 2013). This effect is not limited to depressive disorders; people with generalized anxiety disorder who read a biological description of the etiology of the disorder felt decreased personal responsibility for their symptoms, but also displayed increased prognostic pessimism (Lebowitz, Pyun, & Ahn, 2014). In the present studies, we examined whether people with a mental illness responded differently to essentialist explanations, even when an underlying biological cause was not mentioned. It is possible that people with a mental illness have greater knowledge about mental illness and experience with treatment than the general population, so they may interpret essentialist information differently. Having greater knowledge about treatment may make someone more resistant to the negative effects of essentialist framings of mental illness.

Present studies

In this article, we present three studies examining whether essentialism is causally related to the perceived effectiveness of different treatments and to people's stigmatization toward individuals with mental illness. In the present studies, we presented adults with different vignettes about an individual with an artificial mental illness to examine how essentialist beliefs influence their views on treatment effectiveness and stigmatization. We used artificial mental illnesses to more easily manipulate participants' perception of the illness (either essentialist or not). We manipulated the vignettes by emphasizing essentialist-consistent, essentialist-inconsistent, or neutral aspects of the mental illnesses. Critically, none of the vignettes made explicit reference to biology or internal causes, to isolate the effect of essentialist framing on people's beliefs about treatment efficacy and stigmatization.

Rather than asking participants whether they have been diagnosed with a specific disorder (e.g., depression), we simply asked participants whether they have even been diagnosed with a disorder. As participants were judging novel disorders, we did not believe that any specific diagnosis would be more informative than another. We hypothesized that when participants read essentialist-consistent vignettes they would perceive drug treatment to be more effective and talk therapy to be less effective than when reading essentialist-inconsistent descriptions. We also predicted that people with a mental illness would believe that both drug treatment and talk therapy would be more effective than people without a mental illness. In addition, we

hypothesized that highlighting essentialist-consistent aspects would increase stigmatization, as shown by participants' greater desire to distance themselves socially from the individual with the disorder. Finally, we hypothesized that people with a mental illness would have fewer stigmatizing attitudes than people without a mental illness.

STUDY 1

Method

Participants

Participants included 196 adults who completed the study online through Amazon Mechanical Turk (MTurk), an online platform where participants can complete tasks such as participating in research. Twenty-eight participants were removed from analysis because of inattention (failing two attention checks), resulting in a final sample of 168 participants. The final sample included 106 men, 61 women, and one nonbinary gender participants. The mean age was 32.5 years (SD = 7.82). The sample included 98 White/European American, 22 Asian/Asian American, four Hispanic/Latinx, 22 Black/African American, 13 American Indian/Alaska Native, and nine multiracial participants. Forty-five participants reported having a diagnosed mental illness.

Design

We used a two-condition, between-groups design. We presented two vignettes to each participant. The vignette included a description of the disorder followed by a social distancing scale, and questions about perceived treatment efficacy (more details below). The first disorder was always essentialist-neutral. We used this neutral vignette to get a baseline measure of participants' stigmatization and perceived treatment effectiveness. The second disorder could be either essentialist consistent or essentialist inconsistent.

Materials

Vignettes The vignettes were based on descriptions of artificial mental disorders developed by Marsh and Shanks (2014). They were modified to describe a single person using gender-neutral names. Each vignette named an individual, provided a name for the disorder they had, and listed four symptoms of the disorder. The vignettes also included information that emphasized essentialist-consistent, essentialist-inconsistent, or neutral aspects of the disorder. These descriptions were of approximately equal length (see Appendix 1).

Social Distance Scale After viewing each vignette, participants responded to the Social Distance Scale (Link, Phelan, Bresnahan,

Stueve, & Pescosolido, 1999). This scale involves participants rating from 1 (*definitely willing*) to 4 (*definitely unwilling*) how willing they would be to move next door to the person in the vignette, to spend an evening socializing with the person, to make friends with the person, and to have the person marry into the family. This measure showed high internal consistency (α = .85). Social distance scales are widely used to measure stigma (Link, Yang, Phelan, & Collins, 2004).

Perceived treatment efficacy Participants answered two questions about treatment effectiveness: "How effective will drug treatment be at treating the disorder?" and "How effective will talk therapy be at treating the disorder?" Responses were given on a scale from 1 (*extremely ineffective*) to 7 (*extremely effective*).

Essentialist Beliefs Scale Participants completed the Essentialist Beliefs Scale (EBS; Haslam, Rothschild, & Ernst, 2000) for each vignette. This scale consists of nine questions assessing beliefs about individual subconstructs of essentialism. We modified the scale to include examples to increase the clarity of the questions. Responses were on a scale from 1 to 9, with some items reverse coded (see Appendix 2). This scale serves as a manipulation check, to examine whether our descriptions in the vignettes influenced participants' beliefs about each disorder as intended.

Social Desirability Scale Participants completed the Reynolds (1982) Short Form C, one of the most widely used versions of the Marlowe–Crowne Social Desirability Scale. Short Form C includes nine questions that examine whether participants may be untruthfully responding to provide more socially desirable answers. Responses were in a true–false format, with some of the items reverse coded (see Appendix 3). This scale was included to examine whether or not social desirability played a role in participants' responses to the Social Distance Scale.

Procedure

Participants viewed two vignettes. We randomized the order of the disorders, such that each disorder was equally likely to appear first or second. The first vignette was always essentialist-neutral (control). The second was either an essentialist-consistent or essentialist-inconsistent vignette, depending on the condition the participant was randomly assigned to. After each vignette, participants completed the Social Distance Scale, the questions about treatment efficacy, and the Essentialist Beliefs Scale. After reading and responding to both vignettes, participants completed the Social Desirability Scale and a demographics section that included whether or not the participant had ever been diagnosed with a mental illness.

Results

We used four separate general linear models to analyze participants' EBS scores, perceived drug effectiveness, perceived therapy effectiveness, and stigma scores. We included framing condition, mental illness diagnosis, baseline measure (e.g., EBS, drug effectiveness, therapy effectiveness, or stigma for the first disorder, where appropriate), and social desirability as predictors. We also included an interaction between essentialist framing and mental illness diagnosis. We first present the results for the EBS, then for perceived therapy effectiveness, and finally stigma.

EBS

We used the EBS as a manipulation check. As hypothesized, we found that participants who saw the essentialist-consistent framing had higher EBS scores (showing more essentialist reasoning; $M = 6.04$, $SD = 0.84$) than participants who saw the essentialist-inconsistent framing ($M = 5.08$, $SD = 0.85$), $t(153) = 7.29$, $p < .001$. This suggests that our manipulation worked as intended, and participants who read the essentialist-consistent framing essentialized the disorder more than those who read the essentialist-inconsistent framing. We also found an effect of baseline EBS, such that those that had higher EBS scores (i.e., greater essentialist beliefs about categories in general) at baseline still had higher scores after reading the essentialist-consistent or essentialist-inconsistent framing, $t(153) = 3.11$, $p = .002$. We did not find any other effects or interactions, including those for mental illness diagnosis.

Drug therapy effectiveness

As hypothesized, participants who saw the disorder with the essentialist-consistent framing thought that drug treatment would be more effective ($M = 4.87$, $SD = 1.19$) than participants who saw the disorder with the essentialist-inconsistent framing did ($M = 4.35$, $SD = 1.30$), $t(153) = 3.75$, $p < .001$. Contrary to our hypothesis, there was no effect of mental illness diagnosis, $t(153) = -0.95$, $p = .343$. There was an effect of baseline drug effectiveness, such that participants who thought drug treatment was effective at baseline still thought it would be effective after the manipulation, $t(153) = 5.15$, $p < .001$. There was an effect of stigma, such that participants with high stigma scores thought that drug therapy would be less effective than participants with low stigma scores. $t(153) = -2.23$, $p = .027$. No other effects or interactions were significant, including mental illness diagnosis (see Fig. 1).

Talk therapy effectiveness

Contrary to our hypothesis, participants who saw the disorder with the essentialist-consistent framing did not differ in perceived effectiveness of talk therapy from participants who saw the

disorder with the essentialist-inconsistent framing, $t(153) = 0.10$, $p = .917$. As hypothesized, there was an effect of mental illness diagnosis, in that individuals with a mental illness diagnosis thought talk therapy was more effective ($M = 4.47$, $SD = 1.52$) than did those without a diagnosis ($M = 4.32$, $SD = 1.40$), $t(153) = 2.03$, $p = .044$. There was an effect of baseline therapy effectiveness, such that participants who thought talk therapy was effective at baseline still thought it would be effective after the manipulation, $t(153) = 5.79$, $p < .001$. No other effects or interactions were significant (see Fig. 1).

Stigma

Contrary to hypothesis, reading an essentialist-consistent or essentialist-inconsistent framing for a mental illness did not lead to differences in stigma scores, $t(154) = 0.47$, $p = .639$. As hypothesized, we found a main effect of mental illness diagnosis, $t(154) = -4.09$, $p < .001$. Individuals with a mental illness diagnosis had lower stigma scores ($M = 1.95$, $SD = 0.67$) than did those without a diagnosis ($M = 2.59$, $SD = 0.70$). We also found an effect of baseline stigma, such that participants with high stigma at baseline also had high stigma scores after the manipulation, $t(154) = 7.06$, $p < .001$. No other effects or interactions were significant (see Fig. 1).

Discussion

The results of this study suggest that essentialist framing influences people's views on the effectiveness of drug therapy. However, we did not find evidence that essentialist framing influenced perceived talk therapy effectiveness or stigma. This lack of effects is not due to a weak manipulation of essentialist beliefs as the EBS showed that our manipulation did influence how participants essentialized the disorders. In addition, people diagnosed with a mental illness reported lower levels of stigma toward the person in the vignette. It is possible that people who have been diagnosed with a mental illness are more compassionate toward other individuals with mental illnesses because of their own personal experiences with mental illness and stigmatization.

Participants who read an essentialist-consistent vignette believed drug therapy would be more effective than did participants who read an essentialist-inconsistent vignette. This might be because the essentialist framing suggested that there was a single, discrete biological cause behind the mental illness (even though no cause was explicitly mentioned). Participants might believe that a treatment option that addresses underlying biology might have been more effective. People who read the essentialist-inconsistent framing were not cued toward an underlying biological cause and thus did not perceive the drug treatment to be as effective. This result supports the link between essentialist models of mental illness and treatment choices.
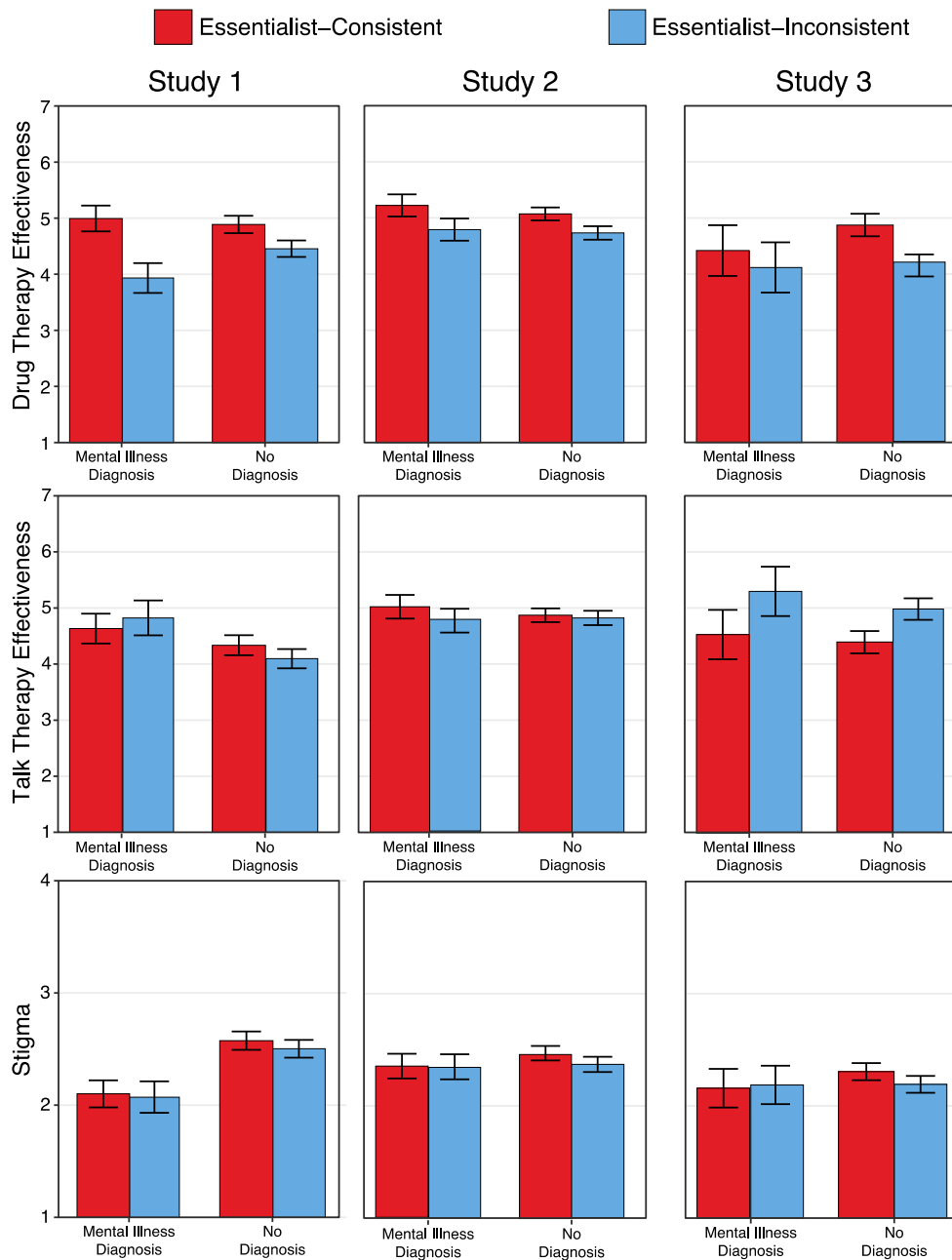
Fig. 1 Participant judgements for drug therapy effectiveness (top panel), talk therapy effectiveness (middle panel), and stigma (bottom panel). The x-axis shows whether or not participants reported having a mental illness diagnosis. The left column presents the results for Study 1 ($n = 168$), the middle column presents the results for the first manipulated disorder of Study 2 ($n = 246$), and the right column presented the results for the first manipulated disorder of Study 3 ($n = 103$). Error bars represent the between-subject standard error of the point estimate

We also found that people who had been diagnosed with a mental illness thought that talk therapy would be more effective than did people who had not been diagnosed with a mental illness. This could be due to positive personal experiences with talk therapy or might reflect a more hopeful view toward treatment in general.

Presenting people with an essentialist framing seemed to lead them to think of the mental illness in an essentialist manner. Some research suggests that interventions, such as the use of generic language, that lead people to essentialize categories still have an effect even after they are explicitly contradicted (Foster-Hanson, Leslie, & Rhodes, 2019). In our study, this would be the equivalent of showing participants first an essentialist-consistent framing followed by an essentialist-inconsistent framing. Therefore, in Study 2, we randomly assigned participants to see first an essentialist-consistent or essentialist-inconsistent framing followed by the opposite framing. This design allows us to examine whether people's

essentialist beliefs about mental illness can be changed in the moment depending on which framing they encounter or whether the first framing shaped their reasoning throughout the study (even if they receive conflicting information later on).

STUDY 2

In this study, participants first saw a disorder with neutral framing, and then saw a disorder with either the essentialist-consistent or the essentialist-inconsistent framing. This study serves as a replication of Study 1. Then, participants saw a disorder with the opposite framing and a disorder with neutral framing to examine whether participants' beliefs change whether they see a contradictory framing.

Method

Participants

There were 306 participants in Study 2, recruited online through Amazon Mechanical Turk. Sixty participants were removed from analysis because they failed attention checks built into the survey, resulting in a final sample of 246 participants. This final sample included 167 men and 77 women participants (two individuals did not respond). The mean age was 31.27 years ($SD = 7.64$). The sample included 138 White/European American, 39 Asian/Asian American, 11 Hispanic/Latinx, 43 Black/African American, four American Indian/Alaska Native, and 11 multiracial participants. Sixty-one participants reported having a diagnosed mental illness.

Design

We used a pre–post intervention design, with framing as a within-groups variable and order of the framings as a between groups variable. We presented four disorders to each participant. As in Study 1, the first disorder was always essentialist-neutral to tap into participants' baseline stigma and perceived treatment effectiveness. For the second disorder, participants were randomly assigned to see an essentialist-consistent or an essentialist-inconsistent framing. For the third disorder, participants saw the opposite framing. The final disorder was essentialist-neutral and was used to see whether there were lasting effects.

Materials

All materials were identical to Study 1, except that participants saw four vignettes instead of two.

Procedure

Participants viewed four vignettes in Study 2. The first and last were always neutral vignettes. The second and third were either essentialist-consistent or essentialist-inconsistent vignettes. The order in which the disorders were presented was randomized across participants. After each vignette, participants completed the Social Distance Scale, the questions about treatment efficacy, and the Essentialist Beliefs Scale. After reading and responding to all vignettes, participants completed the Social Desirability Scale and a demographics section.

Results

We used a similar analytic approach as used in Study 1. We included the baseline measures in all the models for later vignettes. We first present the results for the EBS, then for perceived drug effectiveness, followed by perceived therapy effectiveness, and finally stigma. For each outcome measure, we present a replication of the findings in Study 1 and an extension of these results. For the replication of Study 1, we analyze participants' responses to the second disorder (which is the first time they encounter the framings). For the extension, we analyze their responses to the third and fourth disorders. When looking at the fourth disorder (which had the same framing for all participants), we examined whether there were any differences between those who first saw the essentialist-consistent framing or the essentialist-inconsistent framing.

EBS

Study 1 replication As in Study 1, we found that our manipulation worked as intended, with participants who read the essentialist-consistent framing having higher EBS scores ($M = 5.85$, $SD = 0.73$) than participants who read the essentialist-inconsistent framing ($M = 5.34$, $SD = 0.87$), $t(225) = 5.16$, $p < .001$. We also found that those with higher baseline EBS still had high EBS scores after the manipulation, $t(225) = 6.12$, $p < .001$. No other effects or interactions were significant, including the effect of mental illness diagnosis.

Extension When participants read the opposite framing, we found that those who read the essentialist-consistent framing had higher EBS scores ($M = 5.86$, $SD = 0.77$) than those who read the essentialist-inconsistent framing ($M = 5.24$, $SD = 0.81$), $t(225) = 6.12$, $p < .001$. For the last disorder (when there was no difference in framing), we did not find an effect of which essentialist framing participants saw first on EBS scores, $t(225) = -1.16$, $p = .248$. This suggests that our framing did influence people's essentialist beliefs in the moment. Additionally, we found that baseline EBS predicted EBS scores for the third and fourth disorders, $t(225) = 5.74$, $p <$

.001, and $t(225) = 8.40$, $p < .001$ respectively. No other effects were significant, including the effect of mental illness diagnosis.

Drug therapy effectiveness

Study 1 replication As in the previous study, participants who read the disorder with the essentialist-consistent framing thought that drug treatment would be more effective ($M = 5.10$, $SD = 1.14$) than participants who read the disorder with the essentialist-inconsistent framing ($M = 4.76$, $SD = 1.35$), $t(225) = 2.40$, $p = .017$. We also replicated the effect of baseline drug effectiveness, such that individuals who thought drug treatments were effective at baseline still thought they were effective after the manipulation, $t(225) = 7.84$, $p < .001$. As in Study 1, we found that higher stigma was related to lower perceived effectiveness of drug therapy, $t(225) = -3.17$, $p = .002$. We also found an effect of social desirability, such that higher social desirability was related to higher perceived effectiveness of drug therapy, $t(225) = 3.07$, $p = .002$. There was no effect of mental illness diagnosis (see Fig. 1).

Extension Surprisingly, when participants read the opposite framing, we did not find an effect of essentialist framing, $t(225) = -0.93$, $p = .353$. There was also no effect in the final disorder (when there was no difference in framing), $t(225) = 1.12$, $p = .262$. We found the same effect of baseline drug effectiveness such that participants who thought drug treatments were effective at baseline still thought they would be effective for the third and fourth disorder, $t(225) = 9.32$, $p < .001$, and $t(225) = 8.77$, $p < .001$. We also found the same effect of social desirability for the third disorder, $t(219) = 2.59$, $p = .010$. No other effects or interactions were significant, including the effect of mental illness diagnosis.

Given that we did not find an effect of our manipulation after the first exposure, we decided to conduct an exploratory analysis. It could be that once participants receive the first manipulation, they set their beliefs about the effectiveness of drug therapies for the remainder of the study, but that the effects weaken over time. This would mean that the first manipulation has an effect on the perceived effectiveness of drug therapy, and the effect gets smaller with subsequent framings. To test this hypothesis, we fitted a linear mixed-effects model predicting the perceived effectiveness of drug therapy from whether participants saw the essentialist-consistent (coded 0.5) or the essentialist-inconsistent (coded −0.5) framing first, trial (mean-centered), the interaction between first condition and trial, mental illness diagnosis, perceived effectiveness of drug treatment at baseline, and social desirability. We also included a by-subject random intercept and a by-subject random slope for the effect of trial (and allowed them to correlate). We used a Kenward–Rogers approximation to calculate the degrees of freedom.

We found an overall effect of condition, such that participants who saw the essentialist-consistent framing first perceived drug therapy as being more effective through the remainder of the study than did those who read the essentialist-inconsistent framing first, $F(1, 227) = 5.34$, $p = .022$. Although Fig. 2 shows that this effect weakens over time, we did not find an Initial Condition × Trial interaction, $F(1, 230) = 1.44$, $p = .231$. There was also no overall effect of trial, $F(1, 230) = 0.81$, $p = .368$. We found the same effect of baseline drug effectiveness, $F(1, 227) = 151.42$, $p < .001$, and social desirability, $F(1, 227) = 11.06$, $p = .001$. We did not find an effect of mental illness diagnosis, $F(1, 227) = 2.59$, $p = .109$.

Talk therapy effectiveness

Study 1 replication As in Study 1, we did not find an effect of essentialist framing on talk therapy effectiveness, $t(225) = 0.86$, $p = .393$. As in the previous study, we found that participants who thought talk therapy was effective at baseline still thought it was effective after the manipulation, $t(225) = 10.42$, $p < .001$. There was also an effect of stigma, $t(225) = -3.96$, $p < .001$. No other effects were significant, including the effect of mental illness diagnosis (see Fig. 1).

Extension We did not find an effect of essentialist framing on talk therapy effectiveness for the third disorder, $t(225) = 0.23$, $p = .082$. There was also no effect in the final disorder (when there was no difference in framing), $t(225) = -0.32$, $p = .751$. We found an effect of baseline talk therapy effectiveness such
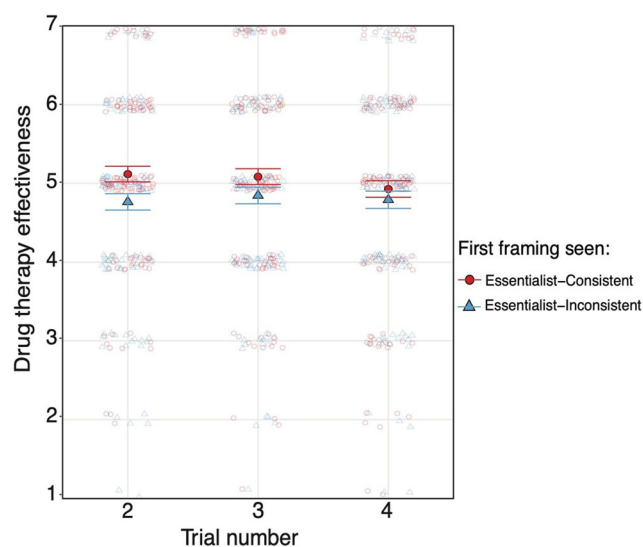

Fig. 2 Model predictions showing the effect of condition on drug effectiveness for each disorder (i.e., trial). The second disorder was the first time that participants received the manipulation. In the third disorders, participants received the opposite manipulation. In the fourth disorder, participants received no manipulation. Error bars represent the within-subjects standard error of the point estimate

that participants who thought that talk therapy was effective at baseline still thought it would be effective for the third and fourth disorder, $t(225) = 9.84$, $p < .001$, and $t(225) = 7.40$, $p < .001$ respectively. We also saw an effect of stigma for both the third and fourth disorders, $t(225) = -3.00$, $p = .003$, and $t(225) = -3.60$, $p < .001$, respectively. No other effects were significant, including the effects of mental illness diagnosis.

Stigma

Study 1 replication As in Study 1, we did not find an effect of essentialist framing, $t(226) = 0.58$, $p = .560$. We also found that those with higher baseline stigma still had high levels of stigma after the manipulation, $t(226) = 12.46$, $p < .001$. No other effects were significant, including the effect of mental illness diagnosis (see Fig. 1).

Extension We did not find an effect of essentialist framing on stigma for either the third disorder, $t(226) = 1.51$, $p = .133$, or fourth disorder, $t(226) = 1.34$, $p = .180$. We found the same effect of stigma for the third and fourth disorder, $t(226) = 13.98$, $p < .001$, and $t(226) = 15.21$, $p < .001$, respectively. No other effects were significant, including the effect of mental illness diagnosis.

Discussion

This study replicates the finding that using an essentialist framing for mental illnesses leads participants to believe that drug therapy will be more effective. However, this was only the case the first time participants read about the disorder. We did not find this relation for stigma or perceived effectiveness of talk therapy. We did not find that any of the effects depended on whether participants had a mental illness.

There was no effect of framing when participants were exposed to the opposite framing in the third disorder. This is surprising because our analysis of the EBS scores suggests that the framing did lead to differences in participants' essentialist beliefs. Our exploratory analysis suggests that participants were influenced by whichever framing they saw first, as those who saw the essentialist-consistent framing first still thought drug therapy was more effective after encountering the opposite framing. The effects of the essentialist-consistent framing degraded over time. The fact that only the first framing was effective suggests that maybe repeated exposure (even to the same framing) does not make a difference. We examine this possibility in Study 3.

STUDY 3

To test whether repeated exposure to the same framing produces the same effects or whether the framing is only relevant on the first exposure, in Study 3 we used a between-groups design so that participants did not see both essentialist-consistent and essentialist-inconsistent information in the vignettes. In Study 2, the framing that was first presented to participants seemed to influence how they responded to the subsequent vignettes. This might be because the within-groups design highlighted the essentialist language, or it could be due to the first manipulated vignette shaping how participants think about mental illness for the duration of the study. Given the order effects that occurred in Study 2, Study 3 used a between-participants design to examine the order effect more closely.

Method

Participants

Participants included 111 undergraduate students enrolled in an introductory psychology course at a large Midwestern university who participated for extra credit in the course. Eight participants were removed from analysis because they failed attention checks in the survey, resulting in a final sample of 103 participants. This final sample included 39 men and 64 women. The mean age was 19.1 years ($SD = 1.53$). The sample included 63 White/European American, 28 Asian/Asian American, five Hispanic/Latinx, three Black/African American, and four multiracial participants. Sixteen participants reported having a mental illness diagnosis.

Design

We used a pre–post intervention design with a between-groups manipulation. Participants saw four disorders. The first disorder served as baseline and was always essentialist-neutral. Participants were randomly assigned to see either essentialist-consistent or essentialist-inconsistent framings for the second and third disorder. The final disorder was also essentialist-neutral.

Materials

Vignettes, Social Distance Scale, and perceived treatment efficacy questions were identical to those in Studies 1 and 2. However, to simplify our analysis, we did not include the EBS or the Social Desirability Scale.

Procedure

Participants viewed a total of four vignettes in Study 3. Participants all viewed two neutral vignettes, and either two essentialist-consistent or two essentialist-inconsistent vignettes depending on the condition they were randomly assigned to. The first and last vignettes presented were always

essentialist-neutral vignettes. The second and third vignettes presented were always either essentialist-consistent or essentialist-inconsistent vignettes. After each vignette, participants completed the Social Distance Scale and answered questions about treatment effectiveness. After reading and responding to all four vignettes, participants answered demographic questions.

Results

We used the same analytic approach as in the previous studies, except that we did not include social desirability in the models, as participants did not complete that measure in this study. For each outcome, we first present the result for the first time that participants see the framing, as this is a replication of Study 1. Then, we present the extension of the findings. For the replication of Study 1, we analyze participants' responses to the second disorder (which is the first time they encounter the framings). For the extension, we analyze their responses to the third and fourth disorders. When looking at the effects of framing on the fourth disorder (which had the same framing for all participants), we looked at the effect of the framing they saw on the previous two disorders.

Drug therapy effectiveness

Replication Surprisingly, we did not find that participants who read the essentialist-consistent framing thought that drug treatment would be more effective ($M = 4.83$, $SD = 1.08$) than did participants who read the essentialist-inconsistent framing ($M = 4.08$, $SD = 1.47$), $t(92) = 1.46$, $p = .146$. However, the means were in the direction consistent with the findings from the other studies. We replicated the effect of baseline drug effectiveness, such that participants who thought drug treatment was effective at baseline still thought that drug treatment was effective after the manipulation, $t(92) = 2.89$, $p = .005$. No other effects were significant, including the effect of mental illness diagnosis (see Fig. 1).

Extension We did not find any evidence that the framing manipulation had any effect when participants read it again in the third disorder, $t(92) = 0.75$, $p = .455$, or when they saw no framing in the final disorder, $t(92) = 0.49$, $p = .626$. We found the same effect of baseline drug effectiveness for the third disorder, $t(92) = 2.38$, $p = .020$, but not for the fourth disorder, $t(92) = 1.88$, $p = .063$. For the fourth disorder, we found that participants who had a mental illness diagnosis thought drug therapy was more effective ($M = 4.43$, $SD = 1.20$) than did those without a diagnosis ($M = 5.19$, $SD = 1.17$), $t(92) = 2.06$, $p = .042$. No other effects were significant.

Once again, we examined whether the effect of the initial framing decreased over time. To test this effect, we fitted a linear mixed-effects model predicting the perceived

effectiveness of drug therapy from condition (coded −0.5 for essentialist-inconsistent and 0.5 for essentialist-consistent), trial (mean-centered), the interaction between condition and trial, mental illness diagnosis, and perceived effectiveness of drug at baseline. We also included a by-subject random intercept and a by-subject random slope for the effect of trial (and allowed the two to correlate). We used a Kenward–Rogers approximation to calculate the degrees of freedom. As in Study 2, we found an effect of framing across all trials, such that participants who saw the essentialist-consistent framing continued to perceive drug therapy as being more effective through the remainder of the study than did those who read the essentialist-inconsistent framing, $F(1, 94.01) = 5.00$, $p = .028$. As can be seen in Fig. 3, once again, we did not find a Condition × Trial interaction, $F(1, 96) = 2.07$, $p = .154$. There was also no overall effect of trial, $F(1, 96) = 0.35$, $p = .555$. We found the same effect of baseline drug effectiveness, $F(1, 94) = 10.72$, $p = .001$. We did not find an effect of mental illness diagnosis, $F(1, 94) = 2.07$, $p = .785$.

Talk therapy effectiveness

Replication In line with our initial hypothesis, but contrary to the other previous studies, we found that participants who read the essentialist-consistent framing perceived talk therapy as less effective ($M = 4.37$, $SD = 1.37$) than participants who read the essentialist-inconsistent framing did ($M = 5.12$, $SD = 1.26$), $t(92) = -2.01$, $p = .048$. We replicated the effect of baseline therapy effectiveness, such that participants that thought talk therapy was effective at baseline still thought it was effective after the manipulation, $t(92) = 4.29$, $p < .001$. No other effects or interactions were significant, including the effect of mental illness diagnosis (see Fig. 1).

Extension We did not find an effect of essentialist framing on the perceived effectiveness of talk therapy for either the third or fourth disorders, $t(92) = -1.02$, $p = .309$, and $t(92) = -1.00$, $p = .318$, respectively. We found an effect of baseline therapy effectiveness for the third disorder, $t(92) = 3.46$, $p = .001$, but not the fourth, $t(92) = 1.97$, $p = .051$. No other effects were significant, including the effect of mental illness diagnosis.

Given that we found an effect of framing on the first disorder, we conducted an exploratory analysis to see whether the effect weakened over time. We fitted a linear mixed-effects model predicting the perceived effectiveness of talk therapy from condition (coded −0.5 for essentialist-inconsistent and 0.5 for essentialist-consistent), trial (mean-centered), the interaction between condition and trial, mental illness diagnosis, and perceived effectiveness of talk therapy at baseline. We also included a by-subject random intercept and a by-subject random slope for the effect of trial (and allowed the two to correlate). We used a Kenward–Rogers approximation to calculate the degrees of freedom. We did not find an overall
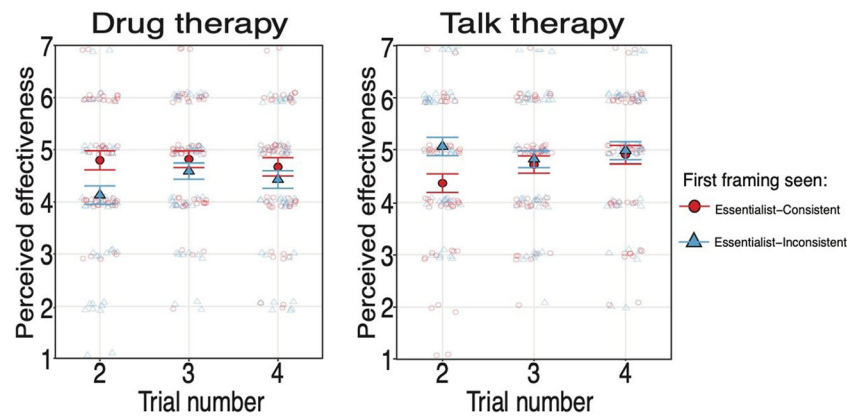
Fig. 3 Model predictions showing the effect of condition on drug effectiveness (left panel) and talk therapy effectiveness (right panel) for each disorder (i.e., trial). In the fourth disorder participants received no manipulation. Error bars represent the within-subjects standard error of the point estimate

effect of framing across all trials, $F(1, 94) = 2.67$, $p = .106$, or trial, $F(1, 96) = 2.66$, $p = .106$. However, we found a Condition × Trial interaction, $F(1, 96) = 4.83$, $p = .030$. As can be seen in Fig. 3, reading the essentialist-consistent framing led participants to perceive talk therapy as being less effective than did those who read the essentialist-inconsistent framing, but this effect rapidly disappeared over time. We found the same effect of baseline talk therapy effectiveness, $F(1, 94) = 18.48$, $p < .001$. We did not find an effect of mental illness diagnosis, $F(1, 94) = 0.95$, $p = .331$.

Stigma

Replication As in the previous studies, we did not find an effect of essentialist framing, $t(93) = 0.31$, $p = .754$. We did find an effect of baseline stigma, such that those with high levels of stigma at baseline still had high levels after the manipulation, $t(93) = 7.29$, $p < .001$. No other effects were significant, including the effect of mental illness diagnosis (see Fig. 1).

Extension We did not find an effect of essentialist framing on stigma for the third or fourth disorders, $t(93) = 0.22$, $p = .826$, and $t(93) = 0.37$, $p = .715$, respectively. We did find the same effect of baseline stigma for both disorders, $t(93) = 8.58$, $p < .001$, and $t(93) = 6.66$, $p < .001$, for the third and fourth disorder, respectively. No other effects were significant, including the effect of mental illness diagnosis.

Discussion

We did not find that the essentialist-consistent framing significantly increased participants' perceived effectiveness of drug therapy when they first encounter it. However, we did find that, across all the trials, those who saw the essentialist-consistent framing thought that drug therapy was more effective than those who saw the essentialist-inconsistent framing.

Additionally, participants who read the essentialist-consistent framing thought that talk therapy would be less effective than did people who read the essentialist-inconsistent framing. Participants might have thought that talk therapy would be less effective, as it does not alter any internal structures. However, given that this is the only study in which we found an effect for the effectiveness of talk therapy, and this study has the smallest sample size, this effect might not be reliable. This study, along with the results of Study 2, suggest that framing mental illnesses in an essentialist manner has an effect when participants first encounter it and then the effect fades over time. In line with the previous two studies, this study suggests that essentialist language does not affect stigma (as measured by social distancing).

Combined analysis

It is possible that we did not find an interaction between framing and mental illness diagnosis because of the low number of participants with a mental illness diagnosis in each study. Additionally, we wanted to check whether finding the predicted effect of framing on talk therapy was spurious or whether the effect was small, and so we found it only once. In this section, we combine data from all three studies to test whether this is the case. Combining all participants means we have a sample of 368 participants without a mental illness diagnosis and 122 participants with a diagnosis. We first examine whether our two populations had differences at baseline. Then, we analyze the data from the second disorder (where participants first encounter the manipulation) to see whether our results change with more power.

Baseline disorder

We ran three general linear models—one predicting drug therapy effectiveness, one predicting talk therapy effectiveness, and one predicting stigma. We used population

(undergraduate students coded as −0.5 and MTurk workers coded as 0.5), mental illness diagnosis, and their interaction. We found that MTurk workers ($M = 4.76$, $SD = 1.25$) thought that drug therapy would be more effective than did undergraduates ($M = 4.34$, $SD = 1.21$), $t(486) = 2.08$, $p = .038$. People who reported having a mental illness diagnosis ($M = 5.02$, $SD = 1.19$) thought that drug therapy would be more effective than those who reported not having a diagnosis ($M = 4.55$, $SD = 1.28$), $t(486) = 2.38$, $p = .018$. There were no differences between our two populations in perceived effectiveness of talk therapy, $t(486) = −0.60$, $p = .545$, or stigma, $t(486) = 0.73$, $p = .465$. There was also no effect of mental illness diagnosis for either talk therapy effectiveness, $t(486) = 1.65$, $p = .100$, or stigma, $t(486) = −0.69$, $p = .491$. The interaction between population and mental illness diagnosis was not significant in any of the analyses.

Manipulation

We ran three general linear models—one predicting drug therapy effectiveness, one predicting talk therapy effectiveness, and one predicting stigma. We used framing condition, mental illness diagnosis, their interaction, and baseline ratings as predictors.

Drug therapy effectiveness We found that participants who saw the essentialist-consistent framing ($M = 4.97$, $SD = 1.15$) thought that drug therapy would be more effective than did those who read the essentialist-inconsistent framing ($M = 4.49$, $SD = 1.38$), $t(485) = 4.50$, $p < .001$. We did not find an effect of mental illness diagnosis on perceived effectiveness of drug therapy, $t(485) = −0.11$, $p = .909$. There was no interaction, $t(485) = 0.80$, $p = .426$. We found an effect of baseline drug effectiveness, $t(485) = 10.42$, $p < .001$ (see Fig. 4).

Talk therapy effectiveness We did not find an effect of framing on perceived effectiveness of talk therapy, $t(485) = −0.36$, $p = .720$. We did find that participants with a mental illness diagnosis ($M = 4.99$, $SD = 1.46$) thought that talk therapy would be more effective than participants without a diagnosis ($M = 4.56$, $SD = 1.45$), $t(485) = 2.78$, $p = .014$. There was no interaction, $t(485) = 0.12$, $p = .904$. We found an effect of baseline talk therapy effectiveness, $t(485) = 12.98$, $p < .001$ (see Fig. 4).

Stigma We did not find an effect of framing on stigma, $t(485) = 0.71$, $p = .476$. We did find that participants with a mental illness diagnosis ($M = 2.15$, $SD = 0.73$) had lower stigma scores than did participants without a diagnosis ($M = 2.45$, $SD = 0.73$), $t(485) = −3.25$, $p = .001$. There was no interaction, $t(485) = −0.78$, $p = .437$. We found an effect of baseline stigma, $t(485) = 16.36$, $p < .001$ (see Fig. 4).
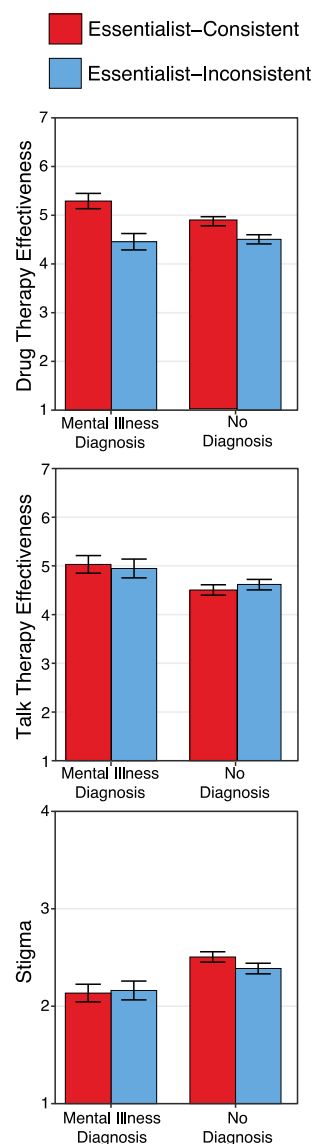

Fig. 4 Participant judgements for drug therapy effectiveness (top panel), talk therapy effectiveness (middle panel), and stigma (bottom panel) for the combined analysis of all studies. The x-axis shows whether or not participants reported having a mental illness diagnosis. Error bars represent the between-subject standard error of the point estimate

Mediation analysis

Given that we found a reliable effect of condition on the perceived effectiveness of drug therapy, we now explore whether this change is in fact due to our manipulation changing participants' essentialist beliefs about mental illness. To do this, we conducted a mediation analysis examining whether the effect of condition on perceived effectiveness of drug therapy is mediated by participants' scores on the EBS. We fit a path model predicting EBS scores from framing condition, and perceived effectiveness from both the EBS and framing condition. We only included 414 participants from Study 1 and Study 2 because participants in Study 3 did not complete

the EBS. We followed the recommendations of Preacher and Hayes (2004) and ran 10,000 simulations, and we tested the indirect effect using nonparametric percentile bootstrapping.

As before, we found that the essentialist-consistent framing led to higher perceived effectiveness of drug therapy ($b = 0.41$, 95% CI [0.17, 0.66]). Additionally, the essentialist-consistent framing led to higher EBS scores ($b = 0.69$, 95% CI [0.53, 0.85]). When controlling for framing condition, higher EBS scores were also related to higher perceived effectiveness of drug therapy ($b = 0.28$, 95% CI [0.13, 0.44]). After controlling for EBS score, there was no effect of the framing condition ($b = 0.13$, 95% CI [−0.04, 0.47]). The indirect effect of framing condition on perceived effectiveness of drug therapy through EBS scores was significant, as the bootstrap confidence interval does not include zero ($b = 0.20$, 95% CI [0.09, 0.32]). This indirect effect represents 48.2% of the total effect of condition on perceived drug effectiveness. Therefore, our data are consistent with the mediational model in which framing had an impact on perceived effectiveness of drugs because it changed participants' essentialist beliefs.

General discussion

Essentialist language played an important role in participants' beliefs about treatment. Collectively, these studies suggest that framing mental illnesses with an essentialist lens increases individuals' essentialist beliefs toward mental illness, which in turn influences their beliefs on drug treatments. We think that the essentialist framing led participants to view mental illnesses as having a distinct internal cause, even when a biological explanation was not explicitly stated. Past research in different domains has suggested that when essentialist language is used, people are more likely to reason using internal causes (Taylor, Rhodes, & Gelman, 2009; Gelman, 2003, 2004). Given that drugs work at a biochemical level, participants might believe that medication is better suited to treat these internal causes. Given that patients' beliefs about treatment might influence treatment adherence and efficacy (Kocsis et al., 2009; Raue et al., 2009), mental health providers should consider this issue when describing treatment options to their patients.

We did not find support for the idea that individuals with a mental illness interpret essentialist information differently than do individuals without a mental illness. It is possible that we did not find the predicted association because we asked whether participants ever received a mental illness diagnosis. Previous studies have focused on participants with a specific mental disorder (Kemp et al., 2014). We did not think this strategy was reasonable for our study, as we presented participants with artificial illnesses. Previous studies suggest that people who belong to a stigmatized group interpret essentialist information differently (Morandini et al., 2017; Morandini

et al., 2015). It is possible that these different interpretations only appear when the information is about your specific group (and not about related groups, especially those defined by a novel or artificial illness).

In addition to essentialist language, views on treatment effectiveness were related to participants' stigma. Participants with higher levels of stigma generally believed both drug and talk therapy to be less effective than did those with lower stigma. Given the correlational nature of our data, we cannot make claims as to the direction of this relation. It may be that people who do not believe that mental illnesses are easily treatable do not want to spend time with people who have a mental illness. Alternatively, people with greater stigma might blame people with a mental illness because they view them as being in control of their symptoms or may be looking for a way to justify their stigmatization. Future research should examine the direction of and mechanisms behind this relationship.

We did not find that essentialist framing influenced stigmatization. This contradicts past research (e.g., Howell et al., 2011) that has demonstrated a relationship between essentialist beliefs and stigma. One potential explanation for these results is that our manipulation was relatively subtle and may not have been strong enough to produce differences in stigmatization. In all of our framings, we described a person as having a diagnosis using person-first language (e.g., "Terry has Mirania") rather than using a noun phrase (e.g., "Terry is a Miraniac"). A preference for using noun phrases to describe someone with a mental illness is associated with holding more essentialist beliefs, as well as greater stigmatization and lower empathy (Howell, Ulan, & Powell, 2014). It has also been found that generic noun phrases (e.g., "Miraniacs behave like this") lead to greater essentializing of categories (Rhodes, Leslie, & Tworek, 2012). However, we did find that even this weak manipulation led to differences in the essentialist beliefs that participants had about the disorders (as measured by the EBS). It is also possible that biological explanations, and not essentialist explanations in general, are related to stigma. Future research should examine how essentialist and biological explanations (independently) influence different components of stigma.

We also did not find that the essentialist framing influenced the perceived effectiveness of talk therapy. We initially hypothesized that reading the essentialist-consistent framing would decrease the perceived effectiveness of talk therapy compared to reading the essentialist-inconsistent framing. We only found this result once, in Study 3, and it was not found in our combined analysis. Our Study 3 sample did not differ from our other samples in their baseline perceived effectiveness of talk therapy, and we statistically controlled for the baseline beliefs and for whether participants had been diagnosed with a mental illness, and therefore differences in baseline beliefs are likely not a reason for the differences in

findings. There could still be important differences between the undergraduate and MTurk sample that led to the differences in results; however, until there is more research on this topic, we consider that our studies suggest there is no effect of essentialist framing on the perceived effectiveness of talk therapy. One possible explanation for this lack of effect could be that the essentialist-consistent framing highlights uniformity among category members, whereas the essentialist-inconsistent framing highlights variation among category members. When reading the essentialist framing, people may view drug treatment as having a single mechanism of action that is likely to be equally effective among all of the, highly similar, category members. However, when reading the essentialist-inconsistent framing, drug treatment would be viewed as less likely to be effective across a wide variety of category members because it only has a single mechanism of action. Conversely, people may view psychotherapy as having multiple mechanisms of action and tailored to the individual, so it does not matter whether category members are highly similar or different. However, this is just speculative, and future work should consider exploring people's intuitive understanding of why psychotherapies are effective.

A limitation of these studies is that findings with artificial disorders might not generalize to real mental illnesses. When thinking about someone with a mental illness, people will likely rely on their previous knowledge about that specific disorder or previous experiences with someone with a similar disorder. Nonetheless, using artificial disorder vignettes provides the benefit of limiting participants' prior knowledge and experience with a disorder, which may influence results and limit the power of the manipulation. Although we acknowledge this is a limitation of this series of studies, it was necessary to try to isolate the effect of other individual differences (e.g., mental illness diagnosis).

We also found that participants' beliefs about the efficacy of a treatment seemed to be shaped with the first framing, such that subsequent framings did not matter. This was the case even when the framings were congruent (i.e., both framings either consistent or inconsistent with essentialist views). It is possible that the first framing people encounter shapes their thinking or primes them to think in a particular way (Foster-Hanson et al., 2019). We did see that the effects faded over time, suggesting that if manipulations were spaced out over a longer interval, we might see an effect of presenting more framings.

An important direction for future research is to examine clinicians' beliefs about mental illness and how these may influence treatment. Clinicians have been shown to hold essentialist beliefs about mental illness (Ahn et al., 2006), and past research has demonstrated that clinicians with biomedical training (i.e., psychiatrists) reported less empathy for their patients than did clinicians with less biomedical training (i.e., psychologists and social workers; Lebowitz & Ahn, 2014). In addition, clinicians reported less empathy for a

hypothetical patient when reading biological explanations for their mental illness than when reading psychosocial explanations (Lebowitz & Ahn, 2014). Reading biological explanations of symptoms led clinicians to perceive medication as more effective and psychotherapy as less effective than reading psychosocial explanations (Lebowitz & Ahn, 2014). Future research should examine whether clinician essentialist beliefs, rather than biological explanations, influence the therapeutic choices and treatment outcome.

Essentialist beliefs affect how people think about mental illness—specifically, how they think about treatment for mental illness. These beliefs can be modified by describing mental illnesses in a manner that is consistent or inconsistent with essentialist beliefs. When people read essentialist-consistent explanations for mental illness, they believe that drug treatment will be more effective than when they read essentialist-inconsistent explanations. Researchers, clinicians, and potentially organizers of antistigma campaigns should carefully consider how they talk about mental illness and should avoid talking about mental illness in essentialist or exclusively biological ways.

Open practices statements The data and materials for all experiments are available (https://osf.io/bt26h/?view_only= e45f9f3da49c412bb11f4aaaee6b85bd), and none of the experiments was preregistered.

Appendix 1

Vignettes

Base descriptions

Terry has a mental disorder called Fortioris. People with this disorder display an intense fear of gaining weight. In addition, they lack remorse when they hurt others. They display repetitive motor mannerisms such as repeatedly slapping their hands together. They often act as if they are "on the go," constantly moving around.

Terry has a mental disorder called Obstatia. People with this disorder are very impulsive and will often do things without thinking them through. They also lose a significant amount of weight without dieting. They lose interest and pleasure in all activities. They also have a lot of muscle tension.

Terry has a mental disorder called Nollentism. People with this disorder find hidden and threatening meanings in nonthreatening conversations. They have reduced feelings for others. They often interrupt others when they are speaking. They exercise more than most people.

Terry has a mental disorder called Mirania. People with this disorder are in a depressed mood most of the day. They rarely need to sleep. They have difficulty concentrating. They also

believe their friends and peers are disloyal, even when there is no reason to believe that they are.

Manipulations

Essentialism-consistent framing: This disorder has discrete boundaries, so that people either have the disorder or they don't. It has been in the *Diagnostic Manual* since the first manual was made, and psychologists around the worlds agree on its existence. The disorder is very uniform, so people with the disorder are highly similar to one another.

Essentialism-inconsistent framing: This disorder exists on a continuum, because the disorder can have varying levels of severity. Psychologists debate its existence, and it seems to be present in some cultures but not others. The disorder is very diverse, so that people with the disorder may be very different from each other.

Control framing: Patients with this disorder often start to exhibit symptoms during adolescence. It is usually diagnosed by a psychologist. Experts have yet to determine what percentage of the population has this disorder.

Appendix 2

Essentialist Beliefs Scale

1. Some categories do not allow their members to belong to other categories. For example, a triangle cannot also be a square. On the other hand, some categories do not limit which other categories their members can belong to. For example, a square is also a rectangle.

   Rate the extent to which you believe members of this category (people with [Mirania/Fortioris/Obstatia/Nollentism]) are excluded from being part of other categories.

2. Some categories have necessary features or characteristics; without these characteristics someone cannot be a category member. For example, all triangles must have three sides. Other categories have many similarities, but no features or characteristics are necessary for membership. For example, there are similarities but no necessary features to define something as art.

   Rate the extent to which this category (people with [Mirania/Fortioris/Obstatia/Nollentism]) has the necessary features (features that members need to have in order to belong to the category).

3. Some categories have an underlying reality; although their members have similarities and differences on the surface, underneath they are basically the same. For example, a butterfly and a caterpillar may look different on the surface, but underneath they are the same species.

Other categories also have similarities and differences on the surface, but do not correspond to an underlying reality. For example, butterflies and moths may look similar on the surface, but are not the same species.

   Rate the extent to which this category (people with [Mirania/Fortioris/Obstatia/Nollentism]) has an underlying reality.

4. Some categories have always existed and their characteristics have not changed much throughout history. For example, rocks have always existed. Other categories are less stable and may not have always existed. For example, computers are a relatively new category.

   Rate the extent to which this category (people with [Mirania/Fortioris/Obstatia/Nollentism]) has been stable across time.

5. Membership in some categories is easy to change. For example, water can turn into ice. For other categories it is difficult for category members to become nonmembers. For example, silver cannot turn into gold.

   Rate the extent to which it easy for members of this category (people with [Mirania/Fortioris/Obstatia/Nollentism]) to change their membership.

6. Some categories, like rubies, are more natural than others, whereas others, like plastic, are more artificial.

   Rate the extent to which this category (people with [Mirania/Fortioris/Obstatia/Nollentism]) is natural or artificial.

7. Some categories allow people to make many judgements about their members. Knowing that something belongs to the category tells us a lot about that category member. For example, knowing that something is a living thing can tell you a lot about it. For other categories, knowledge of membership is not very informative. For example, knowing something is a blue thing doesn't tell you very much about it.

   Rate the extent to which the category (people with [Mirania/Fortioris/Obstatia/Nollentism]) is informative about the characteristics of its members.

8. Some categories contain members who are very similar to one another; they have many things in common. For example, cars are highly similar to one another. Other categories contain members who differ greatly from one another, and don't share many characteristics. For example, different foods may be very different from one another.

   Rate the extent to which members of this category (people with [Mirania/Fortioris/Obstatia/Nollentism]) are diverse or similar.

9. Some categories have sharper boundaries than others. For some, membership is clear-cut, definite, and of an either–or variety; things either belong to the category or they do not. For example, the category of bicycles has clear-cut boundaries. For others, membership is more fuzzy; members belong to the category in varying degrees. For example, the category of big things has no clear boundaries.

Rate the extent to which this category (people with [Mirania/Fortioris/Obstatia/Nollentism]) has clear-cut or fuzzy boundaries.

Appendix 3

Reynolds ([1982](#)) Social Desirability Scale Short Form C

1. It is sometimes hard for me to go on with my work if I am not encouraged.
2. I sometimes feel resentful when I don't get my way.
3. On a few occasions, I have given up doing something because I thought too little of my ability.
4. There have been times when I felt like rebelling against people in authority even though I knew they were right.
5. No matter who I'm talking to, I'm always a good listener.
6. There have been occasions when I took advantage of someone.
7. I'm always willing to admit it when I make a mistake.
8. I sometimes try to get even rather than forgive and forget.
9. I am always courteous, even to people who are disagreeable.
10. I have never been irked when people expressed ideas very different from my own.
11. There have been times when I was quite jealous of the good fortune of others.
12. I am sometimes irritated by people who ask favors of me.
13. I have never deliberately said something that hurt someone's feelings.

References

Ahn, W. K., Flanagan, E. H., Marsh, J. K., & Sanislow, C. A. (2006). Beliefs about essences and the reality of mental disorders. *Psychological Science, 17*(9), 759–766.

Balkcom, E. R., Alogna, V. K., Curtin, E. R., Halberstadt, J. B., & Bering, J. M. (2019). Aversion to organs donated by suicide victims: The role of psychological essentialism. *Cognition, 192*, 104037. https://doi.org/10.1016/j.cognition.2019.104037

Breheny, M. (2007). Genetic attribution for schizophrenia, depression, and skin cancer: Impact on social distance. *New Zealand Journal of Psychology, 36*(3), 154.

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin, 137*(5), 800.

Foster-Hanson, E., Leslie, S. J., & Rhodes, M. (2019). Speaking of kinds: How generic language shapes the development of category representations. *PsyArXiv.* 10.31234/osf.io/28qf7

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought.* Oxford, England: Oxford University Press.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*(9), 404-409.

Gelman, S. A., & Bloom, P. (2000). Young children are sensitive to how an object was created when deciding what to name it. *Cognition, 76*(2), 91–103.

Goldstein, B., & Rosselli, F. (2003). Etiological paradigms of depression: The relationship between perceived causes, empowerment, treatment preferences, and stigma. *Journal of Mental Health, 12*(6), 551–563.

Haslam, N., & Ernst, D. (2002). Essentialist beliefs about mental disorders. *Journal of Social and Clinical Psychology, 21*(6), 628–644.

Haslam, N., & Kvaale, E. P. (2015). Biogenetic explanations of mental disorder: The mixed blessings model. *Current Directions in Psychological Science, 24*(5), 399–404.

Haslam, N., & Levy, S. R. (2006). Essentialist beliefs about homosexuality: Structure and implications for prejudice. *Personality and Social Psychology Bulletin, 32*(4), 471–485.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*(1), 113–127.

Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology, 41*(1), 87–100.

Howell, A. J., Weikum, B. A., & Dyck, H. L. (2011). Psychological essentialism and its association with stigmatization. *Personality and Individual Differences, 50*(1), 95–100.

Howell, A. J., Ulan, J. A., & Powell, R. A. (2014). Essentialist beliefs, stigmatizing attitudes, and low empathy predict greater endorsement of noun labels applied to people with mental disorders. *Personality and Individual Differences, 66*, 33–38.

Kemp, J. J., Lickel, J. J., & Deacon, B. J. (2014). Effects of a chemical imbalance causal explanation on individuals' perceptions of their depressive symptoms. *Behaviour Research and Therapy, 56*, 47–52.

Kim, N. S., & LoSavio, S. T. (2009). Causal explanations affect judgments of the need for psychological treatment. *Judgment and Decision Making, 4*(1), 82.

Kocsis, J. H., Leon, A. C., Markowitz, J. C., Manber, R., Arnow, B., Klein, D. N., & Thase, M. E. (2009). Patient preference as a moderator of outcome for chronic forms of major depressive disorder treated with nefazodone, cognitive behavioral analysis system of psychotherapy, or their combination. *The Journal of Clinical Psychiatry, 70*, 354–361.

Lebowitz, M. S., & Ahn, W. K. (2014). Effects of biological explanations for mental disorders on clinicians' empathy. *Proceedings of the National Academy of Sciences, 111*(50), 17786–17790.

Lebowitz, M. S., & Appelbaum, P. S. (2017). Beneficial and detrimental effects of genetic explanations for addiction. *International Journal of Social Psychiatry, 63*(8), 717–723.

Lebowitz, M. S., Ahn, W. K., & Nolen-Hoeksema, S. (2013). Fixable or fate? Perceptions of the biology of depression. *Journal of Consulting and Clinical Psychology, 81*(3), 518.

Lebowitz, M. S., Pyun, J. J., & Ahn, W. K. (2014). Biological explanations of generalized anxiety disorder: Effects on beliefs about prognosis and responsibility. *Psychiatric Services, 65*(4), 498–503.

Lebowitz, M. S., Rosenthal, J. E., & Ahn, W. K. (2016). Effects of biological versus psychosocial explanations on stigmatization of children with ADHD. *Journal of Attention Disorders, 20*(3), 240–250.

Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology, 27*(1), 363–385.

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*(9), 1328–1333.

Link, B. G., Yang, L. H., Phelan, J. C., & Collins, P. Y. (2004). Measuring mental illness stigma. *Schizophrenia Bulletin, 30*(3), 511–541.

Loughman, A., & Haslam, N. (2018). Neuroscientific explanations and the stigma of mental disorder: A meta-analytic study. *Cognitive Research: Principles and Implications*, *3*(1), 43.

Marsh, J. K., & Romano, A. L. (2016). Lay judgments of mental health treatment options: The mind versus body problem. *MDM Policy & Practice*, *1*(1), 2381468316669361. https://doi.org/10.1177/2381468316669361

Marsh, J. K., & Shanks, L. L. (2014). Thinking you can catch mental illness: How beliefs about membership attainment and category structure influence interactions with mental health category members. *Memory & Cognition*, *42*(7), 1011–1025.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 179 -195). New York: Cambridge University Press.

Morandini, J. S., Blaszczynski, A., Ross, M. W., Costa, D. S., & Dar-Nimrod, I. (2015).Essentialist beliefs, sexual identity uncertainty, internalized homonegativity and psychological well-being in gay men. *Journal of Counseling Psychology*, *62*(3), 413.

Morandini, J. S., Blaszczynski, A., Costa, D. S., Godwin, A., & Dar-Nimrod, I. (2017). Born this way: Sexual orientation beliefs and their correlates in lesbian and bisexual women. *Journal of Counseling Psychology*, *64*(5), 560.

Morton, T. A., & Postmes, T. (2009). When differences become essential: Minority essentialism in response to majority treatment. *Personality and Social Psychology Bulletin*, *35*(5), 656–668.

National Institute of Mental Health. (2017). *Mental illness*. Retrieved from https://www.National Institute of Mental Health.nih.gov/health/statistics/mental-illness.shtml

Phelan, J. C. (2002). Genetic bases of mental illness—A cure for stigma? *Trends in Neurosciences*, *25*(8), 430–431.

Phelan, J. C., Yang, L. H., & Cruz-Rojas, R. (2006). Effects of attributing serious mental illnesses to genetic causes on orientations to treatment. *Psychiatric Services*, *57*(3), 382–387.

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*(4), 717–731.

Raue, P. J., Schulberg, H. C., Heo, M., Klimstra, S., & Bruce, M. L. (2009). Patients' depression treatment preferences and initiation, adherence, and outcome: A randomized primary care study. *Psychiatric Services*, *60*(3), 337–343.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe–Crowne Social Desirability Scale. *Journal of Clinical Psychology*, *38*(1), 119–125.

Rhodes, M., Leslie, S-J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 13526–13531. https://doi.org/10.1073/pnas.1208951109

Roberts, S. O., & Gelman, S. A. (2015). Do children see in black and white? Children's and adults' categorizations of multiracial individuals. *Child Development*, *86*(6), 1830–1847.

Roberts, S. O., Ho, A. K., Rhodes, M., & Gelman, S. A. (2017). Making boundaries great again: Essentialism and support for boundary-enhancing initiatives. *Personality and Social Psychology Bulletin*, *43*(12), 1643–1658.

Rüsch, N., Angermeyer, M. C., & Corrigan, P. W. (2005). Mental illness stigma: Concepts, consequences, and initiatives to reduce stigma. *European Psychiatry*, *20*(8), 529–539.

Schroder, H. S., Dawood, S., Yalch, M. M., Donnellan, M. B., & Moser, J. S. (2015). The role of implicit theories in mental health symptoms, emotion regulation, and hypothetical treatment choices in college students. *Cognitive Therapy and Research*, *39*(2), 120–139. https://doi.org/10.1007/s10608-014-9652-6

Segal, S. P., Kotler, P. L., & Holschuh, J. (1991). Attitudes of sheltered care residents toward others with mental illness. *Psychiatric Services*, *42*(11), 1138–1143.

Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys: Cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development*, *80*, 461–481.

Walker, I., & Read, J. (2002). The differential effectiveness of psychosocial and biogenetic causal explanations in reducing negative attitudes toward "mental illness". *Psychiatry: Interpersonal and Biological Processes*, *65*(4), 313–325.

Waxman, S., Medin, D., & Ross, N. (2007). Folk biological reasoning from a cross-cultural developmental perspective: Early essentialist notions are shaped by cultural beliefs. *Developmental Psychology*, *43*(2), 294.

Yopchick, J. E., & Kim, N. S. (2009). The influence of causal information on judgments of treatment efficacy. *Memory & Cognition*, *37*(1), 29–41.