UC Santa Cruz

Title

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

Authors

Peters, Danielle

Menendez, David

Rosengren, Karl

1

2

3

4

5

6

7   Reframing mental illness:

8   The role of essentialism on stigmatization and perceived treatment efficacy

9

10   Danielle E. Peters

11   David Menendez

12   University of Wisconsin - Madison

13

14   Karl S. Rosengren

15   University of Rochester

16

17

18   Corresponding author: dmenendez@wisc.edu

19

24

29

30

31   Abstract

32   People believe that treatments for illnesses are effective when they target the cause of the illness.

33   Prior work suggests that biological essentialist explanations of mental illness lead people to prefer

34   medications or other pharmacological treatments. However, prior work has not distinguished

35   between biological and essentialist explanations. In three studies (total n = 517), we presented

36   adults with vignettes about an individual with an artificial mental illness and manipulated the

37   descriptions to emphasize or de-emphasize essentialist characteristics. Critically, none of the

38   vignettes made reference to a biological basis for the disorder. Participants rated their willingness

39   to interact with the person described in the vignettes and how effective they believed drug

40   treatment and talk therapy would be on the mental illness. Across the three studies, describing

41   mental illness with an essentialist framing led participants to think drug treatments would be more

42   effective, but there was no effect for stigma or perceived effectiveness of talk therapy. This effect

43   appears to be mediated by how much participants essentialized individuals with the disorder. The

44   first framing that participants encountered seemed to shape their reasoning for the remainder of

45   the study, even if they saw conflicting framing later on. The framing manipulation had similar

46   effects for individuals with and without a mental illness. Results suggest that it is important to

47   consider how mental illness is framed to the general public as it might impact people's treatment

48   preferences.

49

50   Keywords: Cognitive Processes, Essentialism, Social Cognition, Perceived treat Effectiveness

51   Stigma

52

53   Reframing mental illness:

54   The role of essentialism on perceived treatment efficacy and stigmatization

55   One in five Americans suffer from a mental illness (NIMH, 2017), but less than half of

56   the people who had experienced a mental illness in the past year received treatment (NIMH,

57   2017). Patient preferences play a significant role in treatment. Past research has shown that

58   patients with depression who receive treatment that aligns with their own preference for

59   medication or psychotherapy are more likely to initiate and adhere to treatment (Raue,

60   Schulberg, Heo, Klimstra, & Bruce, 2009) and have higher remission rates and lower levels of

61   depression (Kocsis et al., 2009). Given this association between preference and adherence to

62   treatment, it is critical to understand why people might prefer certain treatments or believe them

63   to be more effective. One possible factor that could influence beliefs and attitudes about

64   treatments might be essentialist beliefs.

65   **Essentialism and Mental Illness**

66   Psychological essentialism refers to the notion that people believe that categories have an

67   underlying property (an essence) that determines category membership (Medin & Ortony, 1989).

68   People tend to hold essentialist beliefs about simple categories such as "dogs" or "vegetables,"

69   but also about complex social categories such as race, gender, and sexual orientation (Dar-

70   Nimrod & Heine, 2011). Essentialist beliefs encompass a variety of sub-components. People

71   who hold essentialist beliefs about a category are likely to believe that knowing that someone is a

72   member of a certain category is greatly informative (informativeness), that the category has

73   existed across history with very few changes (historical invariance), that the category has all-or-

74   none boundaries (discreteness), and that members of the category are highly similar to one

75   another (uniformity; Gelman, 2003, 2004; Haslam & Ernst, 2002). They are also likely to believe

76   that it is difficult or impossible for a category member to lose membership (immutability), that

77   there are certain characteristics necessary to be a category member (necessary features), that

78   category membership is due to an inherent underlying reality (inherence), and that the category is

79   naturally occurring rather than socially constructed (naturalness; Gelman, 2003; Haslam & Ernst,

80   2002).

81   Many people hold essentialist beliefs about mental illness, believing that the disorder is

82   caused by something inside the person that is core to their identity (Ahn, Flanagan, Marsh, &

83   Sanislow, 2006). This might be problematic as essentialist beliefs of mental illness affect both

84   stigmatization and views on treatment (Dar-Nimrod & Heine, 2011). One reason why people

85   might hold essentialist beliefs about mental illness is that explanations of mental illness

86   frequently attribute internal biological factors as the causes. This is not unique to mental illness

87   as in other domains people often appeal to an underlying biological cause such as the heart, other

88   organs, DNA, or blood to explain category membership (Balkcom, Alogna, Curtin, Halberstadt,

89   & Bering, 2019; Roberts & Gelman, 2015; Waxman, Medin, & Ross, 2007). Biological framing

90   of mental illness may also serve to activate essentialistist reasoning and lead to negative

91   outcomes such as stigma (Loughman & Haslam, 2018).

92   **Biology, Essentialism, and Perceived Treatment Efficacy**

93   Biological explanations of mental illness influence people's beliefs about the efficacy of

94   treatment in general. Lebowitz, Rosenthal, and Ahn (2012) found that, when reading vignettes

95   about children with ADHD, reading a biological explanation of the disorder decreased stigma but

96   increased doubt about treatment efficacy. Additionally, Marsh and Romano (2016) found that

97   people often think that drug treatment would be more effective for symptoms perceived as

98   medically-based and that talk therapy would be more effective for symptoms perceived as

99   psychological in nature. Beliefs about the etiology of mental illness influence beliefs about what

100   kind of treatment is appropriate. That is, if the person believes that symptoms are due to an

101   internal, biological cause, they might perceive treatments that modify something inside the

102   person to be more effective. Yopchick and Kim (2009) found that when making judgements

103   about treatment efficacy, people consider the root cause of the mental illness to be most

104   important. If the root cause of the illness was described as biological, people believed that drug

105   therapy would be more effective, and if the root cause was described as psychological, they

106   believed psychotherapy would be more effective at treating the illness. Similarly, Lebowitz and

107   Appelbaum (2017) found that reading genetic explanations of addiction increased confidence in

108   pharmacotherapy and decreased confidence in psychotherapy. In addition, Phelan, Yang, and

109   Cruz-Rojas (2006) found that belief in a biological cause of mental illness was related to greater

110   endorsement of hospitalization and medication, but lower expectations that a mental health

111   professional could help treat the illness.

112   Although biological explanations of mental illness frequently attribute internal factors as

113   the cause of the illness, this need not be the case as individuals can hold essentialist beliefs about

114   non-biological categories (e.g., art; Gelman & Bloom, 2000). Therefore, it could be that

115   appealing to internal causes, without mentioning a biological cause, might promote an

116   essentialist view of mental illness. Appealing to an internal cause might also influence treatment

117   decisions (Kim & LoSavio, 2009). For example, Schroeder, Dawood, Yalch, Donnellan, and

118   Moser (2015) found that people who had an essentialist view were more likely to prefer

119   medication alone over psychotherapy or a combined treatment. If this is true, then describing

120   mental illnesses in essentialist terms might lead people to prefer treatments (such as medication)

121   that have effects internally, even if a biological cause is not mentioned in the explanation.

122   **Biology, Essentialism and Stigma**

123   In addition to influencing beliefs about treatment, essentialist beliefs about mental illness

124   might also play a role in stigmatization (Dar-Nimrod & Heine, 2011). Stigmatization is the act of

125   distinguishing and labeling differences between humans, normally with a negative connotation,

126   and separating "us" and "them" based on those social differences (Link & Phelan, 2001). The

127   consequences of stigmatization are wide-ranging, including loss of self-esteem, job

128   discrimination, and avoidance of treatment (Rüsch, Angermeyer, & Corrigan, 2005).

129   Essentialist views about social categories such as race, gender, and sexual orientation

130   have been shown to be related to prejudice and stereotyping (Dar-Nimrod & Heine, 2011).

131   People who hold essentialist beliefs are more likely to support legislation that enhances

132   boundaries between social groups, and these beliefs can be manipulated by providing

133   information that either confirms or disconfirms the essentialist belief (Roberts, Ho, Rhodes, &

134   Gelman, 2017). Researchers have also found that people who hold essentialist beliefs about

135   mental illness have more stigmatizing attitudes about people with a mental illness (Howell,

136   Weikum, & Dyck, 2011). However, some studies have failed to find an association between

137   different sub-components of essentialism and stigmatization of individuals with mental illness

138   (Marsh & Shanks, 2014).

139   Biological explanations of mental illness and essentialist beliefs may interact in complex

140   ways to influence stigmatization. Biological explanations of mental illness may reduce stigma by

141   shifting the blame from the individual to biological factors outside of the individual's control.

142   Indeed, Goldstein and Rosselli (2003) found that people who believe that depression is caused by

143   biological factors were less likely to blame people with depression for their illness. However,

144   biological explanations may increase the stigma towards people with a mental illness (Phelan,

145 2002). For example, Walker and Read (2002) found that people who heard a biological

146 explanation of schizophrenia believed that people with schizophrenia were more dangerous and

147 unpredictable than those who heard a psychosocial (non-essentialist) explanation. Even in the

148 same studies there have been conflicting findings. For example, Breheny (2007) found that

149 providing a genetic explanation for schizophrenia decreased stigmatization, but that providing a

150 genetic explanation for depression increased stigmatization. These results suggest that the

151 relation between stigma and biological or essentialist explanations is complex, as biological

152 essentialist explanations could decrease some components of stigma (such as blame), but

153 increase others (such as social distancing; Haslam & Kvaale, 2015). Currently it is difficult to

154 predict when one of the outcomes will occur. One issue with the prior research is that many of

155 these studies combine essentialist and biological information. This combination makes it difficult

156 to pin point whether different types of information have either beneficial or negative effects. In

157 our study we hope to shed light on this relation by examining the effects of essentialist

158 information on its own.

159 **Essentialism Among People with a Mental Illness**

160   Most of the literature on essentialism and mental illness has focused on the beliefs of the

161 general public, but less attention has been given to how individuals with a mental illness respond

162 to essentialist explanations. People who have received psychiatric services have more accepting

163 attitudes towards others with a mental illness (Segal, Kotler, & Holschuh, 1991; Walker & Read

164 2002), and so might interpret essentialist information differently. It has been found that people

165 who belong to a stigmatized group often respond differently to essentialist information. For

166 example, although essentialist views of sexual orientation (such as "born this way") relate to

167 lower stigma among heterosexual individuals (Haslam & Levy, 2006; Haslam, Rothschild, &

168   Ernst, 2002), essentialist views of sexual orientation have mixed effects among homosexual and

169   bisexual individuals (Morandini, Blaszczynski, Costa, Godwin, & Dar-Nimrod, 2017;

170   Morandini, Blaszczynski, Ross, Costa, & Dar-Nimrod, 2015; Morton & Postmes, 2009).

171   One important context in which people with a mental illness may encounter essentialist

172   information is the context of treatment. Biological explanations of psychological symptoms,

173   when viewed through an essentialist lens, may influence people's beliefs about the course and

174   treatment of their illness. In one study that examined the effects of biological explanations

175   among people with mental illness, Kemp, Lickel, and Deacon (2014) randomly assigned

176   individuals who have had a depressive episode to either being told that the episode was due to a

177   neurochemical imbalance (i.e., a biological explanation) or not. Participants who were told that

178   their depression was caused by a neurochemical imbalance displayed increased perceived stigma,

179   greater pessimism about their diagnosis and treatment, and lower perceived ability to regulate

180   their own negative mood states. These results suggest that biological explanations for mental

181   illness may negatively affect those who suffer from mental illnesses.

182   Similarly, past research has also demonstrated that people with depression who endorse

183   biological explanations for their symptoms display greater prognostic pessimism, the belief that

184   mental illnesses are likely to be stable over time and difficult or impossible to treat (Lebowitz,

185   Ahn, & Nolen-Hoeksema, 2013). This effect is not limited to depressive disorders; people with

186   generalized anxiety disorder who read a biological description of the etiology of the disorder felt

187   decreased personal responsibility for their symptoms but also displayed increased prognostic

188   pessimism (Lebowitz, Pyun, & Ahn, 2014). In the present studies, we examined whether people

189   with a mental illness responded differently to essentialist explanations, even when an underlying

190   biological cause was not mentioned. It is possible that people with a mental illness have greater

191   knowledge about mental illness and experience with treatment than the general population, so

192   they may interpret essentialist information differently. Having greater knowledge about

193   treatment may make someone more resistant to the negative effects of essentialist framings of

194   mental illness.

195   **Present Studies**

196   In this article we present three studies examining whether essentialism is causally related

197   to the perceived effectiveness of different treatments and to people's stigmatization towards

198   individuals with mental illness. In the present studies, we presented adults with different

199   vignettes about an individual with an artificial mental illness to examine how essentialist beliefs

200   influence their views on treatment effectiveness and stigmatization. We used artificial mental

201   illnesses in order to more easily manipulate participants' perception of the illness (either

202   essentialist or not). We manipulated the vignettes by emphasizing essentialist-consistent,

203   essentialist-inconsistent, or neutral aspects of the mental illnesses. Critically, none of the

204   vignettes made explicit reference to biology or internal causes in order to isolate the effect of

205   essentialist framing on people's beliefs about treatment efficacy and stigmatization.

206   Rather than asking participants whether they have been diagnosed with a specific

207   disorder (e.g., depression), we simply asked participants whether they have even been diagnosed

208   with a disorder. As participants were judging novel disorders, we did not believe that any

209   specific diagnosis would be more informative than another. We hypothesized that when

210   participants read essentialist-consistent vignettes they would perceive drug treatment to be more

211   effective and talk therapy to be less effective than when reading essentialist-inconsistent

212   descriptions. We also predicted that people with a mental illness would believe that both drug

213   treatment and talk therapy would be more effective than people without a mental illness. In

214   addition, we hypothesized that highlighting essentialist-consistent aspects would increase

215   stigmatization, as shown by participants' greater desire to distance themselve socially from the

216   individual with the disorder. Finally, we hypothesized that people with a mental illness would

217   have less stigmatizing attitudes than people without a mental illness.

218   **STUDY 1**

219   **Method**

220   **Participants**

221   Participants included 196 adults who completed the study online through Amazon

222   Mechanical Turk (MTurk), an online platform where participants can complete tasks such as

223   participating in research. Twenty-eight participants were removed from analysis due to

224   inattention (failing two attention checks), resulting in a final sample of 168 participants. The

225   final sample included 106 men, 61 women, and 1 nonbinary gender participants. The mean age

226   was 32.5 years ($SD = 7.82$). The sample included 98 White/European American, 22 Asian/Asian

227   American, 4 Hispanic/Latinx, 22 Black/African American, 13 American Indian/Alaska Native,

228   and 9 multiracial participants. Forty-five participants reported having a diagnosed mental illness.

229   **Design**

230   We used a two condition, between groups design. We presented two vignettes to each

231   participant. The vignette included a description of the disorder followed by a social distancing

232   scale, and questions about perceived treatment efficacy (more details below). The first disorder

233   was always essentialist-neutral. We used this neutral vignette to get a baseline measure of

234   participants' stigmatization and perceived treatment effectiveness. The second disorder could be

235   either essentialist-consistent or essentialist-inconsistent.

236   **Materials**

237   **Vignettes.** The vignettes were based on descriptions of artificial mental disorders

238   developed by Marsh and Shanks (2014). They were modified to describe a single person using

239   gender-neutral names. Each vignette named an individual, provided a name for the disorder they

240   had, and listed four symptoms of the disorder. The vignettes also included information that

241   emphasized essentialist-consistent, essentialist-inconsistent, or neutral aspects of the disorder.

242   These descriptions were of approximately equal length (see Appendix A).

243   **Social Distance Scale.** After viewing each vignette, participants responded to the Social

244   Distance Scale (Link, Phelan, Bresnahan, Stueve, & Pescosolido, 1999). This scale involves

245   participants rating from 1 (*definitely willing*) to 4 (*definitely unwilling*) how willing they would

246   be to move next door to the person in the vignette, to spend an evening socializing with the

247   person, to make friends with the person, and to have the person marry into the family. This

248   measure showed high internal consistency ($\alpha = .85$). Social distance scales are widely used to

249   measure stigma (Link, Yang, Phelan, & Collins, 2004).

250   **Perceived Treatment Efficacy.** Participants answered two questions about treatment

251   effectiveness, "How effective will drug treatment be at treating the disorder?" and "How

252   effective will talk therapy be at treating the disorder?" Responses were given on a scale from 1

253   (*extremely ineffective*) to 7 (*extremely effective*).

254   **Essentialist Beliefs Scale.** Participants completed the Essentialist Beliefs Scale (Haslam,

255   Rothschild, & Ernst, 2000) for each vignette. This scale consists of nine questions assessing

256   beliefs about individual sub-constructs of essentialism. We modified the scale to include

257   examples in order to increase the clarity of the questions. Responses were on a scale from 1 – 9,

258   with some items reverse-coded (see Appendix B). This scale serves as a manipulation check, to

259   examine whether our descriptions in the vignettes influenced participants' beliefs about each

260   disorder as intended.

261   **Social Desirability Scale.** Participants completed the Reynolds (1982) Short Form C,

262   one of the most widely used versions of the Marlowe-Crowne Social Desirability Scale. Short

263   Form C includes 9 questions that examine whether participants may be untruthfully responding

264   in order to provide more socially desirable answers. Responses were in a true – false format, with

265   some of the items reverse-coded (see Appendix C). This scale was included to examine whether

266   or not social desirability played a role in participants' responses to the Social Distance Scale.

267   **Procedure**

268   Participants viewed two vignettes. We randomized the order of the disorders, such that

269   each disorder was equally likely to appear first or second. The first vignette was always

270   essentialist-neutral (control). The second was either an essentialist-consistent or essentialist-

271   inconsistent vignette depending on the condition the participant was randomly assigned to. After

272   each vignette, participants completed the Social Distance Scale, the questions about treatment

273   efficacy, and the Essentialist Beliefs Scale. After reading and responding to both vignettes,

274   participants completed the Social Desirability Scale and a demographics section that included

275   whether or not the participant had ever been diagnosed with a mental illness.

276   **Results**

277   We used four separate general linear models to analyze participants' EBS scores,

278   perceived drug effectiveness, perceived therapy effectiveness, and stigma scores. We included

279   framing condition, mental illness diagnosis, baseline measure (e.g., EBS, drug effectiveness,

280   therapy effectiveness, or stigma for the first disorder, where appropriate), and social desirability

281   as predictors. We also included an interaction between essentialist framing and mental illness

282   diagnosis. We first present the results for the EBS, then for perceived therapy effectiveness, and

283   finally stigma.

284   **EBS**

285   We used the EBS as a manipulation check. As hypothesized, we found that participants

286   that saw the essentialist-consistent framing had higher EBS scores (showing more essentialist

287   reasoning; $M = 6.04$, $SD = 0.84$) than participants that saw the essentialist-inconsistent framing

288   ($M = 5.08$, $SD = 0.85$), $t(153) = 7.29$, $p < .001$. This suggests that our manipulation worked as

289   intended and participants that read the essentialist-consistent framing essentialized the disorder

290   more than those who read the essentialist-inconsistent framing. We also found an effect of

291   baseline EBS, such that those that had higher EBS scores (i.e. greater essentialist beliefs about

292   categories in general) at baseline still had higher scores after reading the essentialist-consistent or

293   essentialist-inconsistent framing, $t(153) = 3.11$, $p = .002$. We did not find any other effects or

294   interactions, including those for mental illness diagnosis.

295   **Drug Therapy Effectiveness**

296   As hypothesized, participants who saw the disorder with the essentialist-consistent

297   framing thought that drug treatment would be more effective ($M = 4.87$, $SD = 1.19$) than

298   participants who saw the disorder with the essentialist-inconsistent framing ($M = 4.35$, $SD =$

299   1.30), $t(153) = 3.75$, $p < .001$. Contrary to our hypothesis, there was no effect of mental illness

300   diagnosis, $t(153) = -0.95$, $p = .343$. There was an effect of baseline drug effectiveness, such that

301   participants that thought drug treatment was effective at baseline still thought it would be

302   effective after the manipulation, $t(153) = 5.15$, $p < .001$. There was an effect of stigma, such that

303   participants with high stigma scores thought that drug therapy would be less effective than

304   participants with low stigma scores. $t(153) = -2.23$, $p = .027$. No other effects or interactions

305   were significant, including mental illness diagnosis. See Figure 1.

306   **Talk Therapy Effectiveness**

307   Contrary to our hypothesis, participants who saw the disorder with the essentialist-

308   consistent framing did not differ in perceived effectiveness of talk therapy from participants who

309   saw the disorder with the essentialist-inconsistent framing, $t(153) = 0.10$, $p = .917$. As

310   hypothesized, there was an effect of mental illness diagnosis, in that individuals with a mental

311   illness diagnosis thought talk therapy was more effective ($M = 4.47$, $SD = 1.52$) than those

312   without a diagnosis ($M = 4.32$, $SD = 1.40$), $t(153) = 2.03$, $p = .044$. There was an effect of

313   baseline therapy effectiveness, such that participants that thought talk therapy was effective at

314   baseline still thought it would be effective after the manipulation, $t(153) = 5.79$, $p < .001$. No

315   other effects or interactions were significant. See Figure 1.

316   **Stigma**

317   Contrary to hypothesis, reading an essentialist-consistent or essentialist-inconsistent

318   framing for a mental illness did not lead to differences in stigma scores, $t(154) = 0.47$, $p = .639$.

319   As hypothesized, we found a main effect of mental illness diagnosis, $t(154) = -4.09$, $p < .001$.

320   Individuals with a mental illness diagnosis had lower stigma scores ($M = 1.95$, $SD = 0.67$) than

321   those without a diagnosis ($M = 2.59$, $SD = 0.70$). We also found an effect of baseline stigma,

322   such that participants with high stigma at baseline also had high stigma scores after the

323   manipulation, $t(154) = 7.06$, $p < .001$. No other effects or interactions were significant. See

324   Figure 1.

325
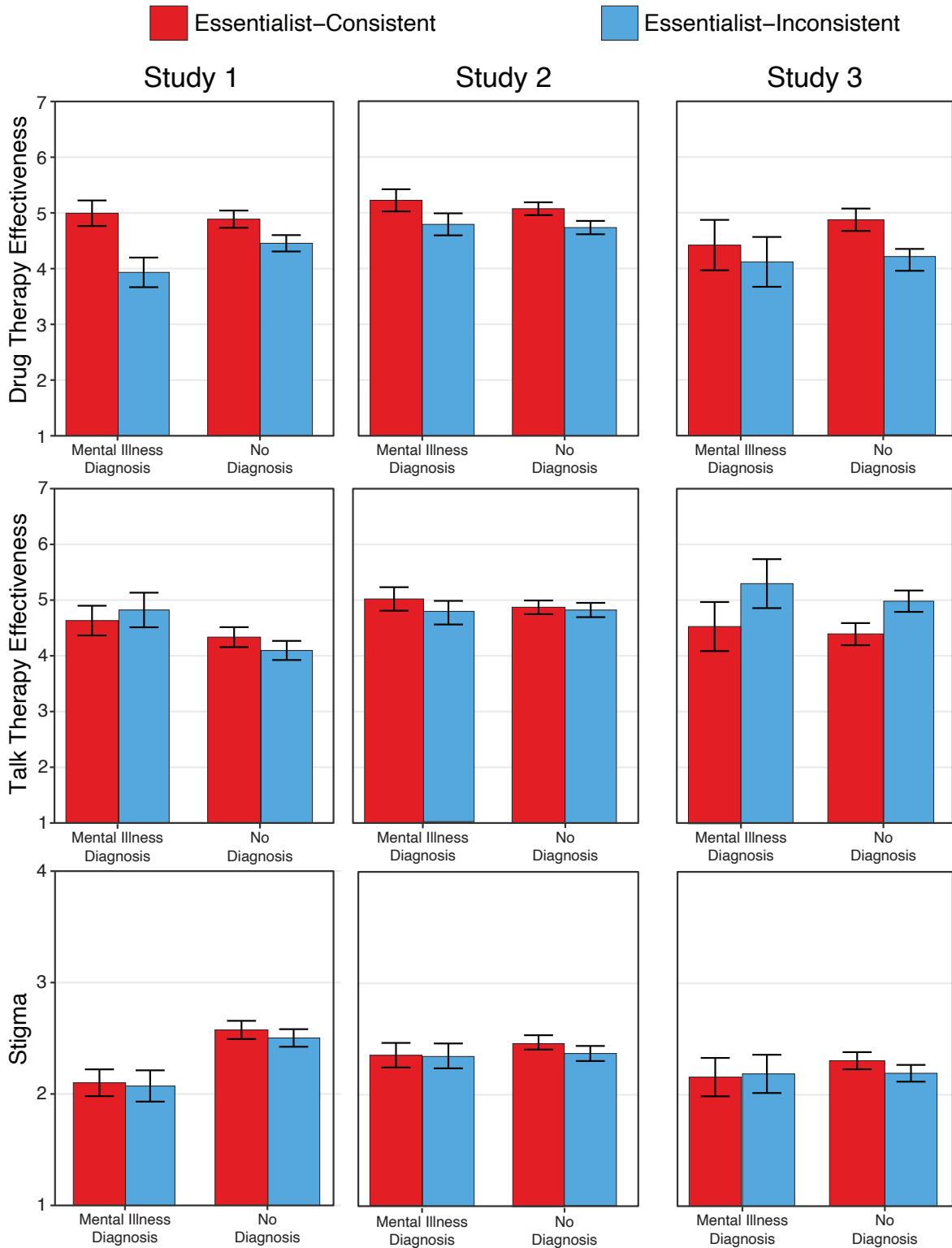
Figure 1. Participant judgements for drug therapy effectiveness (top panel), talk therapy

effectiveness (middle panel), and stigma (bottom panel). The x-axis shows whether or not

328   participants reported having a mental illness diagnosis. The left column presents the results for

329   Study 1 (*n*=168), the middle column presents the results for the first manipulated disorder of

330   Study 2 (*n*=246), and the right column presented the results for the first manipulated disorder of

331   Study 3 (*n*=103). Error bars represent the between-subject standard error of the point estimate.

332

333   **Discussion**

334   The results of this study suggest that essentialist framing influences people's views on the

335   effectiveness of drug therapy. However, we did not find evidence that essentialist framing

336   influenced perceived talk therapy effectiveness or stigma. This lack of effects is not due to a

337   weak manipulation of essentialist beliefs as the EBS showed that our manipulation did influence

338   how participants essentialized the disorders. In addition, people diagnosed with a mental illness

339   reported lower levels of stigma toward the person in the vignette. It is possible that people who

340   have been diagnosed with a mental illness are more compassionate towards other individuals

341   with mental illnesses because of their own personal experiences with mental illness and

342   stigmatization.

343   Participants who read an essentialist-consistent vignette believed drug therapy would be

344   more effective than participants who read an essentialist-inconsistent vignette. This might be

345   because the essentialist framing suggested that there was a single, discrete biological cause

346   behind the mental illness (even though no cause was explicitly mentioned). Participants might

347   believe that a treatment option that addresses underlying biology might have been more

348   effective. People who read the essentialist-inconsistent framing were not cued towards an

349   underlying biological cause and thus did not perceive the drug treatment to be as effective. This

350   result supports the link between essentialist models of mental illness and treatment choices.

351   We also found that people who had been diagnosed with a mental illness thought that talk

352   therapy would be more effective than people who had not been diagnosed with a mental illness.

353   This could be due to positive personal experiences with talk therapy or might reflect a more

354   hopeful view toward treatment in general.

355   Presenting people with an essentialist framing seemed to lead them to think of the mental

356   illness in an essentialist manner. Some research suggests that interventions, such as the use of

357   generic language, that lead people to essentialize categories still have an effect even after they

358   are explicitly contradicted (Foster-Hanson, Leslie, & Rhodes, 2019). In our study, this would be

359   the equivalent of showing participants first an essentialist-consistent framing followed by an

360   essentialist-inconsistent framing. Therefore, in Study 2 we randomly assigned participants to see

361   first an essentialist-consistent or essentialist-inconsistent framing followed by the opposite

362   framing. This design allows us to examine whether people's essentialist beliefs about mental

363   illness can be changed in the moment depending on which framing they encounter or if the first

364   framing shaped their reasoning throughout the study (even if they receive conflicting information

365   later on).

366   **STUDY 2**

367   In this study, participants first saw a disorder with neutral framing, and then saw a

368   disorder with either the essentialist-consistent or the essentialist-inconsistent framing. This study

369   serves as a replication of Study 1. Then, participants saw a disorder with the opposite framing

370   and a disorder with neutral framing to examine whether participants' beliefs change if they see a

371   contradictory framing.

372   **Method**

373   **Participants**

374   There were 306 participants in Study 2, recruited online through Amazon Mechanical

375   Turk. Sixty participants were removed from analysis because they failed attention checks built

376   into the survey, resulting in a final sample of 246 participants. This final sample included 167

377   men and 77 women participants (2 individuals did not respond). The mean age was 31.27 years

378   (SD = 7.64). The sample included 138 White/European American, 39 Asian/Asian American, 11

379   Hispanic/Latinx, 43 Black/African American, 4 American Indian/Alaska Native, and 11

380   multiracial participants. Sixty-one participants reported having a diagnosed mental illness.

381   **Design**

382   We used a pre-intervention-post design with framing as a within-groups variable and

383   order of the framings as a between groups variable. We presented four disorders to each

384   participant. As in Study 1, the first disorder was always essentialist-neutral in order to tap into

385   participants' baseline stigma and perceived treatment effectiveness. For the second disorder,

386   participants were randomly assigned to see an essentialist-consistent or an essentialist-

387   inconsistent framing. For the third disorder, participants saw the opposite framing. The final

388   disorder was essentialist-neutral and was used to see if there were lasting effects.

389   **Materials**

390   All materials were identical to Study 1, except that participants saw four vignettes instead

391   of two.

392   **Procedure**

393   Participants viewed four vignettes in Study 2. The first and last were always neutral

394   vignettes. The second and third were either essentialist-consistent or essentialist-inconsistent

395   vignettes. The order in which the disorders were presented was randomized across participants.

396   After each vignette, participants completed the Social Distance Scale, the questions about

397 treatment efficacy, and the Essentialist Beliefs Scale. After reading and responding to all

398 vignettes, participants completed the Social Desirability Scale and a demographics section.

Results

400 We used a similar analytic approach as used in Study 1. We included the baseline

401 measures in all the models for later vignettes. We first present the results for the EBS, then for

402 perceived drug effectiveness, followed by perceived therapy effectiveness, and finally stigma.

403 For each outcome measure, we present a replication of the findings in Study 1 and an extension

404 of these results. For the replication of Study 1, we analyze participants' responses to the second

405 disorder (which is the first time they encounter the framings). For the extension, we analyze their

406 responses to the third and fourth disorders. When looking at the fourth disorder (which had the

407 same framing for all participants), we examined if there were any differences between those who

408 first saw the essentialist-consistent framing or the essentialist-inconsistent framing.

EBS

410 *Study 1 replication.* As in Study 1, we found that our manipulation worked as intended

411 with participants that read the essentialist-consistent framing having higher EBS scores ($M =$

412 5.85, $SD = 0.73$) than participants who read the essentialist-inconsistent framing ($M = 5.34$, $SD =$

413 0.87), $t(225) = 5.16$, $p < .001$. We also found that those with higher baseline EBS still had high

414 EBS scores after the manipulation, $t(225) = 6.12$, $p < .001$. No other effects or interactions were

415 significant, including the effect of mental illness diagnosis.

416 *Extension.* When participants read the opposite framing, we found that those who read

417 the essentialist-consistent framing had higher EBS scores ($M = 5.86$, $SD = 0.77$) than those who

418 read the essentialist-inconsistent framing ($M = 5.24$, $SD = 0.81$), $t(225) = 6.12$, $p < .001$. For the

419 last disorder (when there was no difference in framing), we did not find an effect of which

420   essentialist framing participants saw first on EBS scores, $t(225) = -1.16$, $p = .248$. This suggests

421   that our framing did influence people's essentialist beliefs in the moment. Additionally, we

422   found that baseline EBS predicted EBS scores for the third and fourth disorders, $t(225) = 5.74$, p

423    $< .001$, and $t(225) = 8.40$, $p < .001$ respectively. No other effects were significant, including the

424   effect of mental illness diagnosis.

425   **Drug Therapy Effectiveness**

426   *Study 1 replication.* As in the previous study, participants who read the disorder with the

427   essentialist-consistent framing thought that drug treatment would be more effective ($M = 5.10$,

428    $SD = 1.14$) than participants who read the disorder with the essentialist-inconsistent framing (M

429    $= 4.76$, $SD = 1.35$), $t(225) = 2.40$, $p = .017$. We also replicated the effect of baseline drug

430   effectiveness, such that individuals who thought drug treatments were effective at baseline still

431   thought they were effective after the manipulation, $t(225) = 7.84$, $p < .001$. As in Study 1, we

432   found that higher stigma was related to lower perceived effectiveness of drug therapy, $t(225) = -$

433   3.17, $p = .002$. We also found an effect of social desirability, such that higher social desirability

434   was related to higher perceived effectiveness of drug therapy, $t(225) = 3.07$, $p = .002$. There was

435   no effect of mental illness diagnosis. See Figure 1.

436   *Extension.* Surprisingly, when participants read the opposite framing, we did not find an

437   effect of essentialist framing, $t(225) = -0.93$, $p = .353$. There was also no effect in the final

438   disorder (when there was no difference in framing), $t(225) = 1.12$, $p = .262$. We found the same

439   effect of baseline drug effectiveness such that participants that thought drug treatments were

440   effective at baseline still thought they would be effective for the third and fourth disorder, $t(225)$

441    $= 9.32$, $p < .001$ and $t(225) = 8.77$, $p < .001$. We also found the same effect of social desirability

442   for the third disorder, $t(219) = 2.59$, $p = .010$. No other effects or interactions were significant,

443   including the effect of mental illness diagnosis.

444   Given that we did not find an effect of our manipulation after the first exposure, we

445   decided to conduct an exploratory analysis. It could be that once participants receive the first

446   manipulation, they set their beliefs about the effectiveness of drug therapies for the remainder of

447   the study, but that the effects weaken over time. This would mean that the first manipulation has

448   an effect on the perceived effectiveness of drug therapy and the effect gets smaller with

449   subsequent framings. To test this hypothesis we fitted a linear mixed-effects model predicting the

450   perceived effectiveness of drug therapy from whether participants saw the essentialist-consistent

451   (coded 0.5) or the essentialist-inconsistent (coded -0.5) framing first, trial (mean-centered), the

452   interaction between first condition and trial, mental illness diagnosis, perceived effectiveness of

453   drug treatment at baseline, and social desirability. We also included a by-subject random

454   intercept and a by-subject random slope for the effect of trial (and allowed them to correlate).

455   We used a Kenward-Rogers approximation to calculate the degrees of freedom.

456   We found an overall effect of condition, such that participants who saw the essentialist-

457   consistent framing first perceived drug therapy as more effective through the remainder of the

458   study than those who read the essentialist-inconsistent framing first, $F(1, 227) = 5.34$, $p = .022$.

459   Although Figure 2 shows that this effect weakens over time, we did not find an initial condition

460   by trial interaction, $F(1, 230) = 1.44$, $p = .231$. There was also no overall effect of trial, $F(1, 230)$

461     $= 0.81$, $p = .368$. We found the same effect of baseline drug effectiveness, $F(1, 227) = 151.42$, p

462     $< .001$, and social desirability, $F(1, 227) = 11.06$, $p = .001$. We did not find an effect of mental

463   illness diagnosis, $F(1, 227) = 2.59$, $p = .109$.
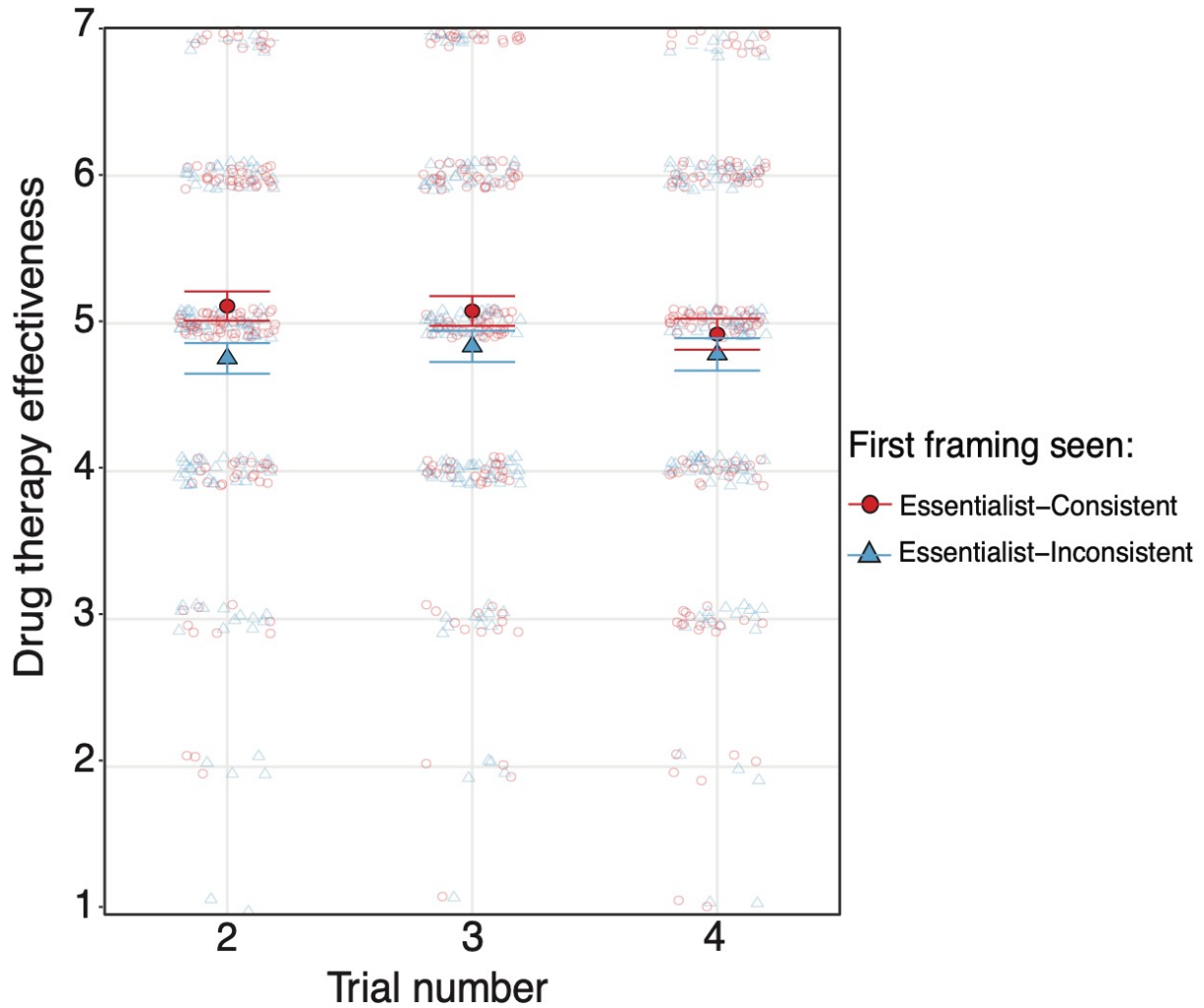
464

465   *Figure 2.* Model predictions showing the effect of condition on drug effectiveness for each

466   disorder (i.e. trial). The second disorder was the first time that participants received the

467   manipulation. In the third disorders participants received the opposite manipulation. In the fourth

468   disorder participants received no manipulation. Error bars represent the within-subject standard

469   error of the point estimate.

470   **Talk Therapy Effectiveness**

471   *Study 1 replication.* As in Study 1, we did not find an effect of essentialist framing on

472   talk therapy effectiveness, $t(225) = 0.86$, $p = .393$. As in the previous study, we found that

473   participants that thought talk therapy was effective at baseline still thought it was effective after

474   the manipulation, $t(225) = 10.42$, $p < .001$. There was also an effect of stigma, $t(225) = -3.96$, $p <$

475    $.001$. No other effects were significant, including the effect of mental illness diagnosis. See

476   Figure 1.

477   *Extension.* We did not find an effect of essentialist framing on talk therapy effectiveness

478   for the third disorder, $t(225) = 0.23$, $p = .082$. There was also no effect in the final disorder

479   (when there was no difference in framing), $t(225) = -0.32$, $p = .751$. We found an effect of

480   baseline talk therapy effectiveness such that participants that thought talk therapy was effective

481   at baseline still thought it would be effective for the third and fourth disorder, $t(225) = 9.84$, $p <$

482    $.001$ and $t(225) = 7.40$, $p < .001$ respectively. We also saw an effect of stigma for both the third

483   and fourth disorders, $t(225) = -3.00$, $p = .003$ and $t(225) = -3.60$, $p < .001$ respectively. No other

484   effects were significant, including the effects of mental illness diagnosis.

485   **Stigma**

486   *Study 1 replication.* As in Study 1, we did not find an effect of essentialist framing, $t(226)$

487    $= 0.58$, $p = .560$. We also found that those with higher baseline stigma still had high levels of

488   stigma after the manipulation, $t(226) = 12.46$, $p < .001$. No other effects were significant,

489   including the effect of mental illness diagnosis. See Figure 1.

490   *Extension.* We did not find an effect of essentialist framing on stigma for either the third

491   disorder, $t(226) = 1.51$, $p = .133$, or fourth disorder, $t(226) = 1.34$, $p = .180$. We found the same

492   effect of stigma for the third and fourth disorder, $t(226) = 13.98$, $p < .001$, and $t(226) = 15.21$, p

493    $< .001$ respectively. No other effects were significant, including the effect of mental illness

494   diagnosis.

495   **Discussion**

496   This study replicates the finding that using an essentialist framing for mental illnesses

497   leads participants to believe that drug therapy will be more effective. However, this was only the

498   case the first time participants read about the disorder. We did not find this relation for stigma or

499   perceived effectiveness of talk therapy. We did not find that any of the effects depended on

500   whether participants had a mental illness.

501   There was no effect of framing when participants were exposed to the opposite framing

502   in the third disorder. This is surprising because our analysis of the EBS scores suggests that the

503   framing did lead to differences in participants' essentialist beliefs. Our exploratory analysis

504   suggests that participants were influenced by whichever framing they saw first, as those who saw

505   the essentialist-consistent framing first still thought drug therapy was more effective after

506   encountering the opposite framing. The effects of the essentialist-consistent framing degraded

507   over time. The fact that only the first framing was effective suggests that maybe repeated

508   exposure (even to the same framing) does not make a difference. We examine this possibility in

509   Study 3.

510   **STUDY 3**

511   To test whether repeated exposure to the same framing produces the same effects or

512   whether the framing is only relevant on the first exposure, in Study 3 we utilized a between-

513   groups design so that participants did not see both essentialist-consistent and essentialist-

514   inconsistent information in the vignettes. In Study 2, the framing that was first presented to

515   participants seemed to influence how they responded to the subsequent vignettes. This might be

516   because the within-groups design highlighted the essentialist language, or it could be due to the

517   first manipulated vignette shaping how participants think about mental illness for the duration of

518   the study. Given the order effects that occurred in Study 2, Study 3 used a between-participants

519   design in order to examine the order effect more closely.

520   **Method**

521   **Participants**

522   Participants included 111 undergraduate students enrolled in an introductory psychology

523   course at a large Mid-Western university who participated for extra-credit in the course. Eight

524   participants were removed from analysis because they failed attention checks in the survey,

525   resulting in a final sample of 103 participants. This final sample included 39 men and 64 women.

526   The mean age was 19.1 years ($SD = 1.53$). The sample included 63 White/European American,

527   28 Asian/Asian American, 5 Hispanic/Latinx, 3 Black/African American, and 4 multiracial

528   participants. Sixteen participants reported having a mental illness diagnosis.

529   **Design**

530   We used a pre-intervention-post design with a between-groups manipulation. Participants

531   saw four disorders. The first disorder served as baseline and was always essentialist-neutral.

532   Participants were randomly assigned to see either essentialist-consistent or essentialist-

533   inconsistent framings for the second and third disorder. The final disorder was also essentialist-

534   neutral.

535   **Materials**

536   Vignettes, Social Distance Scale, and perceived treatment efficacy questions were

537   identical to Studies 1 and 2. However, in order to simplify our analysis, we did not include the

538   EBS or the Social Desirability Scale.

539   **Procedure**

540   Participants viewed a total of four vignettes in Study 3. Participants all viewed two

541   neutral vignettes, and either two essentialist-consistent or two essentialist-inconsistent vignettes

542   depending on the condition they were randomly assigned to. The first and last vignettes

543   presented were always essentialist-neutral vignettes. The second and third vignettes presented

544   were always either essentialist-consistent or essentialist-inconsistent vignettes. After each

545   vignette, participants completed the Social Distance Scale and questions about treatment

546   effectiveness. After reading and responding to all four vignettes, participants answered

547   demographic questions.

548   **Results**

549   We used the same analytic approach as the previous studies, except that we did not

550   include social desirability in the models as participants did not complete that measure in this

551   study. For each outcome we first present the result for the first time participants see the framing

552   as this is a replication of Study 1. Then, we present the extension of the findings. For the

553   replication of Study 1, we analyze participants' responses to the second disorder (which is the

554   first time they encounter the framings). For the extension, we analyze their responses to the third

555   and fourth disorders. When looking at the effects of framing on the fourth disorder (which had

556   the same framing for all participants), we looked at the effect of the framing they saw on the

557   previous two disorders.

558   **Drug Therapy Effectiveness**

559   *Replication.* Surprisingly, we did not find that participants who read the essentialist-

560   consistent framing thought that drug treatment would be more effective ($M = 4.83$, $SD = 1.08$)

561   than participants who read the essentialist-inconsistent framing ($M = 4.08$, $SD = 1.47$), $t(92) =$

562   1.46, $p = .146$. However, the means were in the direction consistent with the findings from the

563   other studies. We replicated the effect of baseline drug effectiveness, such that participants that

564   thought drug treatment was effective at baseline still thought that drug treatment was effective

565   after the manipulation, $t(92) = 2.89$, $p = .005$. No other effects were significant, including the

566   effect of mental illness diagnosis. See Figure 1.

567   *Extension.* We did not find any evidence that the framing manipulation had any effect

568   when participants read it again in the third disorder, $t(92) = 0.75$, $p = .455$, or when they saw no

569   framing in the final disorder, $t(92) = 0.49$, $p = .626$. We found the same effect of baseline drug

570   effectiveness for the third disorder, $t(92) = 2.38$, $p = .020$, but not for the fourth disorder, $t(92) =$

571   1.88, $p = .063$. For the fourth disorder, we found that participants that had a mental illness

572   diagnosis thought drug therapy was more effective ($M = 4.43$, $SD = 1.20$) than those without a

573   diagnosis ($M = 5.19$, $SD = 1.17$), $t(92) = 2.06$, $p = .042$. No other effects were significant.

574   Once again, we examined whether the effect of the initial framing decreased over time.

575   To test this effect we fitted a linear mixed-effects model predicting the perceived effectiveness of

576   drug therapy from condition (coded -0.5 for essentialist-inconsistent and 0.5 for essentialist-

577   consistent), trial (mean-centered), the interaction between condition and trial, mental illness

578   diagnosis, and perceived effectiveness of drug at baseline. We also included a by-subject random

579   intercept and a by-subject random slope for the effect of trial (and allowed the two to correlate).

580   We used a Kenward-Rogers approximation to calculate the degrees of freedom. As in Study 2,

581   we found an effect of framing across all trials, such that participants who saw the essentialist-

582   consistent framing continued to perceive drug therapy as more effective through the remainder of

583   the study than those who read the essentialist-inconsistent framing, $F(1, 94.01) = 5.00$, $p = .028$.

584   As can be seen in Figure 3, once again, we did not find a condition by trial interaction, $F(1, 96) =$

585   2.07, $p = .154$. There was also no overall effect of trial, $F(1, 96) = 0.35$, $p = .555$. We found the

586   same effect of baseline drug effectiveness, $F(1, 94) = 10.72$, $p = .001$. We did not find an effect

587   of mental illness diagnosis, $F(1, 94) = 2.07$, $p = .785$.

Talk Therapy Effectiveness

589   *Replication.* In line with our initial hypothesis but contrary to the other previous studies,

590   we found that participants who read the essentialist-consistent framing perceived talk therapy as

591   less effective ($M = 4.37$, $SD = 1.37$) than participants who read the essentialist-inconsistent

592   framing ($M = 5.12$, $SD = 1.26$), $t(92) = -2.01$, $p = .048$. We replicated the effect of baseline

593   therapy effectiveness, such that participants that thought talk therapy was effective at baseline

594   still thought it was effective after the manipulation, $t(92) = 4.29$, $p < .001$. No other effects or

595   interactions were significant, including the effect of mental illness diagnosis. See Figure 1.

596   *Extension.* We did not find an effect of essentialist framing on the perceived effectiveness

597   of talk therapy for either the third or fourth disorders, $t(92) = -1.02$, $p = .309$ and $t(92) = -1.00$, p

598    $= .318$ respectively. We found an effect of baseline therapy effectiveness for the third disorder,

599    $t(92) = 3.46$, $p = .001$, but not the fourth, $t(92) = 1.97$, $p = .051$. No other effects were significant,

600   including the effect of mental illness diagnosis.

601   Given that we found an effect of framing on the first disorder, we conducted an

602   exploratory analysis to see if the effect weakened over time. We fitted a linear mixed-effects

603   model predicting the perceived effectiveness of talk therapy from condition (coded -0.5 for

604   essentialist-inconsistent and 0.5 for essentialist-consistent), trial (mean-centered), the interaction

605   between condition and trial, mental illness diagnosis, and perceived effectiveness of talk therapy

606   at baseline. We also included a by-subject random intercept and a by-subject random slope for

607   the effect of trial (and allowed the two to correlate). We used a Kenward-Rogers approximation

608   to calculate the degrees of freedom. We did not find an overall effect of framing across all trials,

609     $F(1, 94) = 2.67$, $p = .106$, or trial, $F(1, 96) = 2.66$, $p = .106$. However, we found a condition by

610   trial interaction, $F(1, 96) = 4.83$, $p = .030$. As can be seen in Figure 3, reading the essentialist-

611   consistent framing led participants to perceive talk therapy as less effective than those who read

612   the essentialist-inconsistent framing, but this effect rapidly disappeared over time. We found the

613   same effect of baseline talk therapy effectiveness, $F(1, 94) = 18.48$, $p < .001$. We did not find an

614   effect of mental illness diagnosis, $F(1, 94) = 0.95$, $p = .331$.
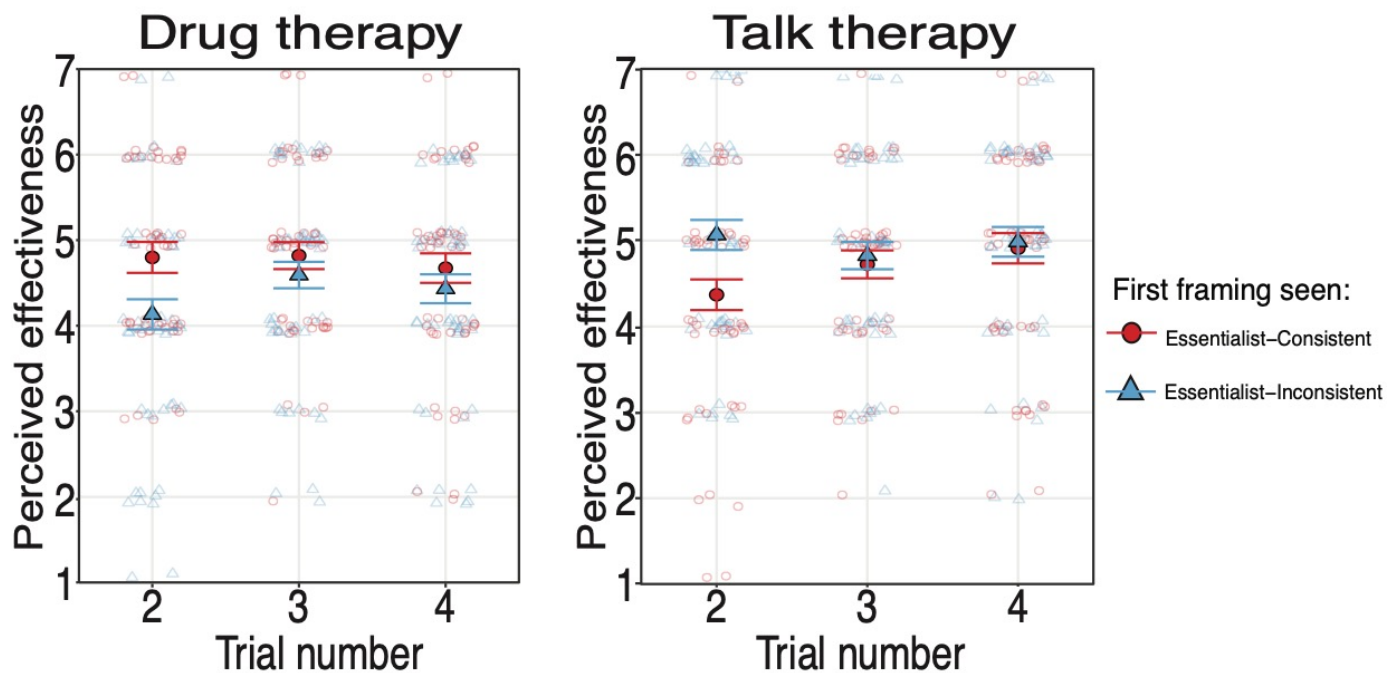
615

616   Figure 3. Model predictions showing the effect of condition on drug effectiveness (left panel)

617   and talk therapy effectiveness (right panel) for each disorder (i.e. trial). In the fourth disorder

618   participants received no manipulation. Error bars represent the within-subject standard error of

619   the point estimate.

620   **Stigma**

621   *Replication.* As in the previous studies, we did not find an effect of essentialist framing,

622     $t(93) = 0.31$, $p = .754$. We did find an effect of baseline stigma, such that those with high levels

623   of stigma at baseline still had high levels after the manipulation, $t(93) = 7.29$, $p < .001$. No other

624   effects were significant, including the effect of mental illness diagnosis. See Figure 1.

625   *Extension.* We did not find an effect of essentialist framing on stigma for the third or

626   fourth disorders, $t(93) = 0.22$, $p = .826$ and $t(93) = 0.37$, $p = .715$ respectively. We did find the

627   same effect of baseline stigma for both disorders, $t(93) = 8.58$, $p < .001$ and $t(93) = 6.66$, $p <$

628     $.001$ for the third and fourth disorder respectively. No other effects were significant, including

629   the effect of mental illness diagnosis.

Discussion

631   We did not find that the essentialist-consistent framing significantly increased

632   participants' perceived effectiveness of drug therapy when they first encounter it. However, we

633   did find that, across all the trials, those who saw the essentialist-consistent framing thought that

634   drug therapy was more effective than those who saw the essentialist-inconsistent framing.

635   Additionally, participants who read the essentialist-consistent framing thought that talk therapy

636   would be less effective than people who read the essentialist-inconsistent framing. Participants

637   might have thought that talk therapy would be less effective as it does not alter any internal

638   structures. However, given that this is the only study in which we found an effect for the

639   effectiveness of talk therapy, and this study has the smallest sample size, this effect might not be

640   reliable. This study, along with the results of Study 2, suggest that framing mental illnesses in an

641   essentialist manner has an effect when participants first encounter it and then the effect fades

642   over time. In line with the previous two studies, this study suggests that essentialist language

643   does not affect stigma (as measured by social distancing).

Combined Analysis

645   It is possible that we did not find an interaction between framing and mental illness

646   diagnosis because of the low number of participants with a mental illness diagnosis in each

647   study. Additionally, we wanted to check whether finding the predicted effect of framing on talk

648   therapy was spurious or if the effect is small, and so we only found it once. In this section, we

649   combine data from all three studies to test if this is the case. Combining all participants means

650   we have a sample of 368 participants without a mental illness diagnosis and 122 participants

651   with a diagnosis. We first examine whether our two populations had differences at baseline,

652   Then, we analyze the data from the second disorder (where participants first encounter the

653   manipulation) to see if our results change with more power.

654   **Baseline Disorder**

655   We ran 3 general linear models— one predicting drug therapy effectiveness, one

656   predicting talk therapy effectiveness, and one predicting stigma. We used population

657   (undergraduate students coded as -0.5 and MTurk workers coded as 0.5), mental illness

658   diagnosis, their interaction. We found that MTurk workers ($M = 4.76$, $SD = 1.25$) thought that

659   drug therapy would be more effective than undergraduates ($M = 4.34$, $SD = 1.21$), $t(486) = 2.08$,

660    $p = .038$. People who reported having a mental illness diagnosis ($M = 5.02$, $SD = 1.19$) thought

661   that drug therapy would be more effective than those who reported not having a diagnosis ($M =$

662   4.55, $SD = 1.28$), $t(486) = 2.38$, $p = .018$. There were no differences between our two populations

663   in perceived effectiveness of talk therapy, $t(486) = -0.60$, $p = .545$, or stigma, $t(486) = 0.73$, $p =$

664    $.465$. There was also no effect of mental illness diagnosis for either talk therapy effectiveness,

665    $t(486) = 1.65$, $p = .100$, or stigma, $t(486) = -0.69$, $p = .491$. The interaction between population

666   and mental illness diagnosis was not significant in any of the analyses.

667   **Manipulation**

668   We ran 3 general linear models— one predicting drug therapy effectiveness, one

669   predicting talk therapy effectiveness, and one predicting stigma. We used framing condition,

670   mental illness diagnosis, their interaction, and baseline ratings as predictors.

671   *Drug Therapy Effectiveness*. We found that participants who saw the essentialist-

672   consistent framing (M = 4.97, SD = 1.15) thought drug therapy would be more effective than

673   those who read the essentialist-inconsistent framing (M = 4.49, SD = 1.38), $t(485) = 4.50$, $p <$

674   .001. We did not find an effect of mental illness diagnosis on perceived effectiveness of drug

675   therapy, $t(485) = -0.11$, $p = .909$. There was no interaction, $t(485) = 0.80$, $p = .426$. We found an

676   effect of baseline drug effectiveness, $t(485) = 10.42$, $p < .001$. See Figure 4.

677   *Talk Therapy Effectiveness*. We did not find an effect of framing on perceived

678   effectiveness of talk therapy, $t(485) = -0.36$, $p = .720$. We did find that participants with a mental

679   illness diagnosis ($M = 4.99$, $SD = 1.46$) thought that talk therapy would be more effective than

680   participants without a diagnosis ($M = 4.56$, $SD = 1.45$), $t(485) = 2.78$, $p = .014$. There was no

681   interaction, $t(485) = 0.12$, $p = .904$. We found an effect of baseline talk therapy effectiveness,

682    $t(485) = 12.98$, $p < .001$. See Figure 4.

683   *Stigma.* We did not find an effect of framing on stigma, $t(485) = 0.71$, $p = .476$. We did

684   find that participants with a mental illness diagnosis ($M = 2.15$, $SD = 0.73$) had lower stigma

685   scores that participants without a diagnosis ($M = 2.45$, $SD = 0.73$), $t(485) = -3.25$, $p = .001$.

686   There was no interaction, $t(485) = -0.78$, $p = .437$. We found an effect of baseline stigma, $t(485)$

687    $= 16.36$, $p < .001$. See Figure 4.
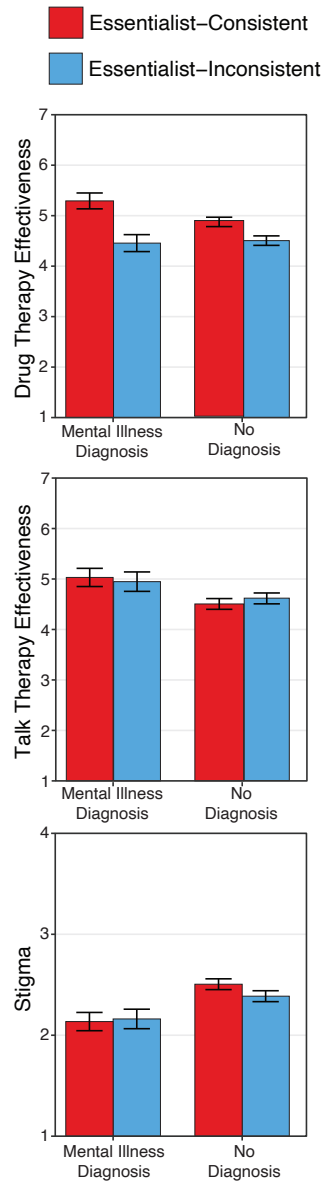
688

689   *Figure 4.* Participant judgements for drug therapy effectiveness (top panel), talk therapy

690   effectiveness (middle panel), and stigma (bottom panel) for the combined analysis of all studies.

691   The x-axis shows whether or not participants reported having a mental illness diagnosis. Error

692   bars represent the between-subject standard error of the point estimate.

693   **Mediation analysis**

694   Given that we found a reliable effect of condition on the perceived effectiveness of drug

695   therapy, we now explore whether this change is in fact due to our manipulation changing

696   participants' essentialist beliefs about mental illness. To do this, we conducted a mediation

697   analysis examining whether the effect of condition on perceived effectiveness of drug therapy is

698   mediated by participants' scores on the EBS. We fit a path model predicting EBS scores from

699   framing condition, and perceived effectiveness from both the EBS and framing condition. We

700   only included 414 participants from Study 1 and 2 because participants in Study 3 did not

701   complete the EBS. We followed the recommendations of Preacher and Hayes (2004) and ran

702   10,000 simulations and we tested the indirect effect using nonparametric percentile

703   bootstrapping.

704   As before, we found that the essentialist-consistent framing led to higher perceived

705   effectiveness of drug therapy, $b = 0.41$, 95% $CI = 0.17, 0.66$. Additionally, the essentialist-

706   consistent framing led to higher EBS scores, $b = 0.69$, 95% $CI = 0.53, 0.85$. When controlling for

707   framing condition, higher EBS scores were also related to higher perceived effectiveness of drug

708   therapy, $b = 0.28$, 95% $CI = 0.13, 0.44$. After controlling for EBS score there was no effect of the

709   framing condition, $b = 0.13$, 95% $CI = -0.04, 0.47$. The indirect effect of framing condition on

710   perceived effectiveness of drug therapy through EBS scores was significant as the bootstrap

711   confidence interval does not include 0, $b = 0.20$, 95% $CI = 0.09, 0.32$. This indirect effect

712   represents 48.2% of the total effect of condition on perceived drug effectiveness. Therefore, our

713   data is consistent with the mediational model in which framing had an impact on perceived

714   effectiveness of drugs because it changed participants' essentialist beliefs.

715   **General Discussion**

716   Essentialist language played an important role in participants' beliefs about treatment.

717   Collectively, these studies suggest that framing mental illnesses with an essentialist lens

718   increases individuals' essentialist beliefs towards mental illness, which in turn influences their

719   beliefs on drug treatments. We think that the essentialist framing led participants to view mental

720   illnesses as having a distinct internal cause, even when a biological explanation was not

721   explicitly stated. Past research in different domains has suggested that when essentialist language

722   is used people are more likely to reason using internal causes (Taylor, Rhodes, & Gelman, 2009;

723   Gelman, 2003; 2004). Given that drugs work at a biochemical level, participants might believe

724   that medication is better suited to treat these internal causes. Given that patients' beliefs about

725   treatment might influence treatment adherence and efficacy (Kocsis et al., 2009; Raue et al.,

726   2009), mental health providers should consider this issue when describing treatment options to

727   their patients.

728   We did not find support for the idea that individuals with a mental illness interpret

729   essentialist information differently than individuals without a mental illness. It is possible that we

730   did not find the predicted association because we asked whether participants ever received any

731   mental illness diagnosis. Previous studies have focused on participants with a specific mental

732   disorder (Kemp et al., 2014). We did not think this strategy was reasonable for our study as we

733   presented participants with artificial illnesses. Previous studies suggest that people who belong to

734   a stigmatized group interpret essentialist information differently (Morandini et al., 2015, 2017).

735   It is possible that these different interpretations only appear when the information is about your

736   specific group (and not about related groups, especially those defined by a novel or artificial

737   illness).

738   In addition to essentialist language, views on treatment effectiveness were related to

739   participants' stigma. Participants with higher levels of stigma generally believed both drug and

740   talk therapy to be less effective than those with lower stigma. Given the correlational nature of

741   our data, we cannot make claims as to the direction of this relation. It may be that people who do

742   not believe mental illnesses are easily treatable do not want to spend time with people who have

743   a mental illness. Alternatively, people with greater stigma might blame people with a mental

744   illness because they view them as in control of their symptoms or may be looking for a way to

745   justify their stigmatization. Future research should examine the direction of and mechanisms

746   behind this relationship.

747   We did not find that essentialist framing influenced stigmatization. This contradicts past

748   research (e.g. Howell et al., 2011) that has demonstrated a relationship between essentialist

749   beliefs and stigma. One potential explanation for these results is that our manipulation was

750   relatively subtle and may not have been strong enough to produce differences in stigmatization.

751   In all of our framings, we described a person as having a diagnosis using person-first language

752   (e.g. "Terry has Mirania") rather than using a noun phrase (e.g. "Terry is a Miraniac"). A

753   preference for using noun phrases to describe someone with a mental illness is associated with

754   holding more essentialist beliefs, as well as greater stigmatization and lower empathy (Howell,

755   Ulan, & Powell, 2014). It has also been found that generic noun-phrases (e.g. "Miraniacs behave

756   like this") lead to greater essentializing of categories (Rhodes, Leslie, & Tworek, 2012).

757   However, we did find that even this weak manipulation led to differences in the essentialist

758   beliefs participants had about the disorders (as measured by the EBS). It is also possible that it is

759   biological explanations, and not essentialist explanations in general, that are related to stigma.

760   Future research should examine how essentialist and biological explanations (independently)

761   influence different components of stigma.

762   We also did not find that the essentialist framing influenced the perceived effectiveness

763   of talk therapy. We initially hypothesized that reading the essentialist-consistent framing would

764   decrease the perceived effectiveness of talk therapy compared to reading the essentialist-

765   inconsistent framing. We only found this result once, in Study 3, and it was not found in our

766   combined analysis. Our Study 3 sample did not differ from our other samples in their baseline

767   perceived effectiveness of talk therapy, and we statistically controlled for the baseline beliefs and

768   for whether participants had been diagnosed with a mental illness, therefore differences in

769   baseline beliefs are likely not a reason for the differences in findings. There could still be

770   important differences between the undergraduate and Mturk sample that led to the differences in

771   results, however, until there is more research on this topic, we consider that our studies suggest

772   there is no effect of essentialist framing on the perceived effectiveness of talk therapy. One

773   possible explanation for this lack of effect could be that the essentialist-consistent framing

774   highlights uniformity among category members, while the essentialist-inconsistent framing

775   highlights variation among category members. When reading the essentialist framing, people

776   may view drug treatment as having a single mechanism of action that is likely to be equally

777   effective among all of the, highly similar, category members. However, when reading the

778   essentialist-inconsistent framing, drug treatment would be viewed as less likely to be effective

779   across a wide variety of category members because it only has a single mechanism of action.

780   Conversely, people may view psychotherapy as having multiple mechanisms of action and

781   tailored to the individual, so it does not matter if category members are highly similar or

782   different. However, this is just speculative, and future work should consider exploring people's

783   intuitive understanding of why psychotherapies are effective.

784   A limitation of these studies is that findings with artificial disorders might not generalize

785   to real mental illnesses. When thinking about someone with a mental illness, people will likely

786   rely on their previous knowledge about that specific disorder or previous experiences with

787   someone with a similar disorder. Nonetheless, using artificial disorder vignettes provides the

788   benefit of limiting participants' prior knowledge and experience with a disorder, which may

789   influence results and limit the power of the manipulation. Although we acknowledge this is a

790   limitation of this series of studies, it was necessary in order to try to isolate the effect of other

791   individual differences (e.g., mental illness diagnosis).

792   We also found that participants' beliefs about the efficacy of a treatment seemed to be

793   shaped with the first framing, such that subsequent framings did not matter. This was the case

794   even when the framings were congruent (i.e., both framings either consistent or inconsistent with

795   essentialist views). It is possible that the first framing people encounter shapes their thinking or

796   primes them to think in a particular way (Foster-Hanson et al., 2019). We did see that the effects

797   faded over time, suggesting that if manipulations were spaced out over a longer interval, we

798   might see an effect of presenting more framings.

799   An important direction for future research is to examine clinicians' beliefs about mental

800   illness and how these may influence treatment. Clinicians have been shown to hold essentialist

801   beliefs about mental illness (Ahn et al., 2006) and past research has demonstrated that clinicians

802   with biomedical training (i.e. psychiatrists) reported less empathy for their patients than

803   clinicians with less biomedical training (i.e., psychologists and social workers; Lebowitz & Ahn,

804   2014). In addition, clinicians reported less empathy for a hypothetical patient when reading

805   biological explanations for their mental illness than when reading psychosocial explanations

806   (Lebowitz & Ahn, 2014). Reading biological explanations of symptoms led clinicians to perceive

807   medication as more effective and psychotherapy as less effective than reading psychosocial

808   explanations (Lebowitz & Ahn, 2014). Future research should examine if clinician essentialist

809   beliefs, rather than biological explanations, influence the therapeutic choices and treatment

810   outcome.

811   Essentialist beliefs affect how people think about mental illness, specifically how they

812   think about treatment for mental illness. These beliefs can be modified by describing mental

813   illnesses in a manner that is consistent or inconsistent with essentialist beliefs. When people read

814   essentialist-consistent explanations for mental illness they believe that drug treatment will be

815   more effective than when they read essentialist-inconsistent explanations. Researchers,

816   clinicians, and potentially organizers of anti-stigma campaigns should carefully consider how

817   they talk about mental illness and should avoid talking about mental illness in essentialist or

818   exclusively biological ways.

819

820   **Open Practices Statements**

821   The data and materials for all experiments are available at

822   https://osf.io/bt26h/?view_only=e45f9f3da49c412bb11f4aaaee6b85bd and none of the

823   experiments were preregistered.

824   References

825   Ahn, W. K., Flanagan, E. H., Marsh, J. K., & Sanislow, C. A. (2006). Beliefs about essences and

826   the reality of mental disorders. *Psychological Science, 17*(9), 759-766.

827   Balkcom, E. R., Alogna, V. K., Curtin, E. R., Halberstadt, J. B., & Bering, J. M. (2019).

828   Aversion to organs donated by suicide victims: The role of psychological essentialism.

829   *Cognition*, *192*, 104037.

830   Breheny, M. (2007). Genetic attribution for schizophrenia, depression, and skin cancer: Impact

831   on social distance. *New Zealand Journal of Psychology*, *36*(3), 154.

832   Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: on the deceptive determinism of

833   DNA. *Psychological Bulletin*, *137*(5), 800.

834   Foster-Hanson, E., Leslie, S.J., & Rhodes, M. (2019). Speaking of kinds: How generic language

835   shapes the development of category representations. PsyArXiv.

836   https://doi.org/10.31234/osf.io/28qf7

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

838   Series in Cognitive Dev.

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

840   404-409.

841   Gelman, S. A., & Bloom, P. (2000). Young children are sensitive to how an object was created

842   when deciding what to name it. *Cognition*, *76*(2), 91-103.

843   Goldstein, B., & Rosselli, F. (2003). Etiological paradigms of depression: The relationship

844   between perceived causes, empowerment, treatment preferences, and stigma. *Journal of*

845   *Mental Health*, *12*(6), 551-563.

846   Haslam, N., & Ernst, D. (2002). Essentialist beliefs about mental disorders. *Journal of Social*

847   *and Clinical Psychology*, *21*(6), 628-644.

848   Haslam, N., & Kvaale, E. P. (2015). Biogenetic explanations of mental disorder: The mixed

849   blessings model. *Current Directions in Psychological Science*, *24*(5), 399-404.

850   Haslam, N., & Levy, S. R. (2006). Essentialist beliefs about homosexuality: Structure and

851   implications for prejudice. *Personality and Social Psychology Bulletin*, *32*(4), 471-485.

852   Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories.

853   *British Journal of Social Psychology*, *39*(1), 113-127.

854   Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with

855   prejudice? *British journal of social psychology, 41*(1), 87-100.

856   Howell, A. J., Ulan, J. A., & Powell, R. A. (2014). Essentialist beliefs, stigmatizing attitudes, and

857   low empathy predict greater endorsement of noun labels applied to people with mental

858   disorders. *Personality and Individual Differences*, *66*, 33-38.

859   Howell, A. J., Weikum, B. A., & Dyck, H. L. (2011). Psychological essentialism and its

860   association with stigmatization. *Personality and Individual Differences, 50*(1), 95-100.

861   Kemp, J. J., Lickel, J. J., & Deacon, B. J. (2014). Effects of a chemical imbalance causal

862   explanation on individuals' perceptions of their depressive symptoms. *Behaviour

863   Research and Therapy*, *56*, 47-52.

864   Kim, N. S., & LoSavio, S. T. (2009). Causal explanations affect judgments of the need for

865   psychological treatment. *Judgment and Decision Making*, *4*(1), 82.

866   Kocsis, J. H., Leon, A. C., Markowitz, J. C., Manber, R., Arnow, B., Klein, D. N., & Thase, M.

867   E. (2009). Patient preference as a moderator of outcome for chronic forms of major

868   depressive disorder treated with nefazodone, cognitive behavioral analysis system of

869   psychotherapy, or their combination. *The Journal of Clinical Psychiatry*, 70, 354-361.

870   Lebowitz, M. S., & Ahn, W. K. (2014). Effects of biological explanations for mental disorders

871   on clinicians' empathy. *Proceedings of the National Academy of Sciences*, *111*(50),

872   17786-17790.

873   Lebowitz, M. S., Ahn, W. K., & Nolen-Hoeksema, S. (2013). Fixable or fate? Perceptions of the

874   biology of depression. *Journal of Consulting and Clinical Psychology*, *81*(3), 518.

875   Lebowitz, M. S., & Appelbaum, P. S. (2017). Beneficial and detrimental effects of genetic

876   explanations for addiction. *International Journal of Social Psychiatry*, *63*(8), 717-723.

877   Lebowitz, M. S., Pyun, J. J., & Ahn, W. K. (2014). Biological explanations of generalized

878   anxiety disorder: effects on beliefs about prognosis and responsibility. *Psychiatric

879   Services*, *65*(4), 498-503.

880   Lebowitz, M. S., Rosenthal, J. E., & Ahn, W. K. (2016). Effects of biological versus

881   psychosocial explanations on stigmatization of children with ADHD. *Journal of Attention

882   Disorders*, *20*(3), 240-250.

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

884   363-385.

885   Link, B. G., Phelan, J. C., Bresnahan, M., Stueve, A., & Pescosolido, B. A. (1999). Public

886   conceptions of mental illness: labels, causes, dangerousness, and social distance.

887   *American Journal of Public Health*, *89*(9), 1328-1333.

888   Link, B. G., Yang, L. H., Phelan, J. C., & Collins, P. Y. (2004). Measuring mental illness stigma.

889   *Schizophrenia Bulletin*, *30*(3), 511-541.

890   Loughman, A., & Haslam, N. (2018). Neuroscientific explanations and the stigma of mental

891   disorder: a meta-analytic study. *Cognitive Research: Principles and Implications*, *3*(1),

892   43.

893   Marsh, J. K., & Romano, A. L. (2016). Lay judgments of mental health treatment options: The

894   mind versus body problem. *MDM policy & practice*, *1*(1), 2381468316669361.

895   Marsh, J. K., & Shanks, L. L. (2014). Thinking you can catch mental illness: How beliefs about

896   membership attainment and category structure influence interactions with mental health

897   category members. *Memory & Cognition*, *42*(7), 1011-1025.

898   Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony

899   (Eds.), *Similarity and Analogical Reasoning* (pp. 179 -195). New York: Cambridge

900   University Press.

901   Morandini, J. S., Blaszczynski, A., Costa, D. S., Godwin, A., & Dar-Nimrod, I. (2017). Born this

902   way: Sexual orientation beliefs and their correlates in lesbian and bisexual women.

903   *Journal of Counseling Psychology*, *64*(5), 560.

904   Morandini, J. S., Blaszczynski, A., Ross, M. W., Costa, D. S., & Dar-Nimrod, I. (2015).

905   Essentialist beliefs, sexual identity uncertainty, internalized homonegativity and

906   psychological wellbeing in gay men. *Journal of Counseling Psychology*, *62*(3), 413.

907   Morton, T. A., & Postmes, T. (2009). When differences become essential: Minority essentialism

908   in response to majority treatment. *Personality and Social Psychology Bulletin*, *35*(5),

909   656-668.

910   National Institute of Mental Health. (2017). Mental Illness. Retrieved from

911   https://www.nimh.nih.gov/health/statistics/mental-illness.shtml

912   Phelan, J. C. (2002). Genetic bases of mental illness–a cure for stigma?. *TRENDS in

913   Neurosciences*, *25*(8), 430-431.

914   Phelan, J. C., Yang, L. H., & Cruz-Rojas, R. (2006). Effects of attributing serious mental

915   illnesses to genetic causes on orientations to treatment. *Psychiatric Services*, *57*(3), 382-

916   387.

917   Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects

918   in simple mediation models. *Behavior Research Methods, Instruments, & Computers,*

919   *36(4),* 717-731.

920   Raue, P. J., Schulberg, H. C., Heo, M., Klimstra, S., & Bruce, M. L. (2009). Patients' depression

921   treatment preferences and initiation, adherence, and outcome: a randomized primary care

922   study. *Psychiatric Services*, *60*(3), 337-343.

923   Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne

924   Social Desirability Scale. *Journal of Clinical Psychology*, *38*(1), 119-125.

925   Rhodes, M., Leslie, S-J., Tworek, C. M. (2012) Cultural transmission of social essentialism.

926   *Proceedings of the National Academy of Science, 109,* 13526-13531.

927   Roberts, S. O., & Gelman, S. A. (2015). Do children see in Black and White? Children's and

928   adults' categorizations of multiracial individuals. *Child Development*, *86*(6), 1830-1847.

929   Roberts, S. O., Ho, A. K., Rhodes, M., & Gelman, S. A. (2017). Making boundaries great again:

930   Essentialism and support for boundary-enhancing initiatives. *Personality and Social*

931   *Psychology Bulletin*, *43*(12), 1643-1658.

932   Rüsch, N., Angermeyer, M. C., & Corrigan, P. W. (2005). Mental illness stigma:

933   concepts,consequences, and initiatives to reduce stigma. *European Psychiatry*, *20*(8),

934   529-539.

935 Schroder, H. S., Dawood, S., Yalch, M. M., Donnellan, M. B., & Moser, J. S. (2015). The role of

936   implicit theories in mental health symptoms, emotion regulation, and hypothetical

937   treatment choices in college students. *Cognitive Therapy and Research*, *39*(2),120-139.

938 Segal, S. P., Kotler, P. L., & Holschuh, J. (1991). Attitudes of sheltered care residents toward

939   others with mental illness. *Psychiatric Services*, *42*(11), 1138-1143.

940 Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys: cows will be cows:

941   Children's essentialist reasoning about gender categories and animal species. *Child*

942   *Development*, *80*, 461-481.

943 Walker, I., & Read, J. (2002). The differential effectiveness of psychosocial and biogenetic

944   causal explanations in reducing negative attitudes toward "mental illness". *Psychiatry:*

945   *Interpersonal and Biological Processes*, *65*(4), 313-325.

946 Waxman, S., Medin, D., & Ross, N. (2007). Folkbiological reasoning from a cross-cultural

947   developmental perspective: early essentialist notions are shaped by cultural beliefs.

948   *Developmental Psychology*, *43*(2), 294.

949 Yopchick, J. E., & Kim, N. S. (2009). The influence of causal information on judgments of

950   treatment efficacy. *Memory & Cognition*, *37*(1), 29-41.