# Essentialist Biases Towards Psychiatric Disorders: Brain Disorders are Presumed Innate

Key words: Dualism, Essentialism, Psychiatric Disorders, Core knowledge

## **Abstract**

A large campaign has sought to destigmatize psychiatric disorders by disseminating the view that they are in fact brain disorders. But when psychiatric disorders are associated with neurobiological correlates, laypeople's attitudes towards patients are harsher, and the prognoses seem poorer (Loughman & Haslam, 2018). Here, we ask whether these misconceptions could result from the essentialist presumption that brain disorders are innate. To this end, we invited laypeople to reason about psychiatric disorders that are diagnosed by either a brain or a behavioral test that were strictly matched for their informative value. Participants viewed disorders as more likely to be innate and immutable when the diagnosis was supported by a brain test as compared to a behavioral test. These results show for the first time that people spontaneously essentialize psychiatric conditions that are linked to the brain, even when the brain probe offers no additional diagnostic or genetic information. This bias suggests that people consider the biological essence of living things as materially embodied.

#### Introduction

In the era of the brain, one would expect the public to treat diseases of the brain on par with conditions that ravage any other part of the body. But surprisingly, mental illness still carries a significant stigma (e.g., Ahn, Kim, & Lebowitz, 2017; Haslam & Kvaale, 2015).

People are not simply oblivious to the fact that mental disorders have biogenetic causes. A large campaign by the US Surgeon General (1999) has sought to inform the public that mental disorders are medical biological conditions (hereafter, the "medicalized" view). But these efforts have partly backfired. While people have become increasingly more aware of the biogenetic origins of psychiatric conditions (Pescosolido et al., 2010; Schomerus et al., 2012), they associate biogenetic and biochemical corelates with poorer prognoses (for meta-analyses: Kvaale, Haslam, & Gottdiener, 2013b; Loughman & Haslam, 2018; for review: Ahn et al., 2017) and they tend to project such conditions to patients' relatives (Bennett, Thirlaway, & Murray, 2008).

These attitudes have been attributed to two principles of intuitive psychology: Essentialism and Dualism (Ahn et al., 2017; Dar-Nimrod & Heine, 2011; Haslam & Kvaale, 2015; Loughman & Haslam, 2018). Together, these principles explain why people shun psychiatric patients, and why they further consider genetic conditions to be as immutable. Recent results, however, hint at the possibility that these intuitive biases extend to brain disorders generally (Loughman & Haslam, 2018)—even in the absence of a known genetic cause. Such attitudes, if they exist, would be puzzling, as it is not immediately clear how they could arise from intuitive psychology.

To address this puzzle, here, we revisit an oft neglected aspect of psychological essentialism—the possibility that biological essence is *embodied* (Haslam, Rothschild, & Ernst, 2000; Lindquist, Gendron, Oosterwijk, & Barrett, 2013; Newman & Keil, 2008). We show how this essentialist reasoning could lead to the presumption that psychiatric conditions that "show up" in the brain (i.e., in the body) are innate. Our experiments test this hypothesis. The results shed light on public attitudes towards mental disorders and the seductive allure of neuroscience (Weisberg, Keil, Goodstein, Rawson, & Gray, 2008). In addition, these findings contribute to the large literature on intuitive essentialism.

#### a. The role of Dualism and Essentialism

Public attitudes towards mental disorders have been attributed to two principles of intuitive psychology—Dualism and Essentialism. Of these two principles, Essentialism is the one most relevant to our present inquiry; Dualism, however, also plays a critical role.

Dualism is the belief that the mind is immaterial and distinct from the body (Bloom, 2004). Per Dualism, it is further the mind, not the body, that is the home of one's psychological core (the true self; Strohminger, Knobe, & Newman, 2017). Additionally, the mind (rather than the body) is credited with one's morality and free will (e.g., Greene & Cohen, 2004; Nichols, 2011).

The "medicalized" view of mental disorders counters these properties of human agency (courtesy of Dualism) by suggesting that a patient's actions arise not from their mind (as required by Dualism) but from their body (Ahn et al., 2017). And indeed, when psychiatric symptoms are attributed to a biological source, they are considered less controllable by the patient (e.g., Deacon & Baird, 2009; Kemp, Lickel, & Deacon, 2014) and less likely to benefit from psychotherapy (Deacon & Baird, 2009; Kim, Ahn, Johnson, & Knobe, 2016; Lebowitz & Ahn, 2014).

By undermining patients' agency, the interaction between the "medicalized" view and Dualism could promote two sets of conflicting social attitudes. On the one hand, the reduction of agency could lead the Dualist to dehumanize the patient (Haslam, 2006), and in so doing, promote negative reactions (for reviews: Ahn et al., 2017; Dar-Nimrod & Heine, 2011; Haslam, 2006). But by placing the responsibility of one's actions on one's body (rather than on one's mind), the "medicalized" view could also prompt the Dualist to reduce blame (Ahn et al., 2017). We note that the reduction in blame could also arise because people consider the patient to not in control of their actions, a view that may not be directly due to Dualism (Haslam & Kvaale, 2015). Either way, social attitudes towards patients are expected to improve. This mixture of negative and positive attitudes is indeed evident in meta-analyses of the literature (Kvaale, Gottdiener, & Haslam, 2013a; Kvaale et al., 2013b; Loughman & Haslam, 2018). Thus, Dualism can partially explain these conflicting social reactions to the "medicalized" view.

To further explain why the "medicalized" view of mental disorders also elicits beliefs in poorer prognoses (for meta-analyses: Kvaale et al., 2013a; Kvaale et al., 2013b; Loughman & Haslam, 2018), the literature has invoked a second principle of intuitive psychology, namely, Essentialism (Dar-Nimrod & Heine, 2011; Haslam & Ernst, 2002; Haslam & Kvaale, 2015).

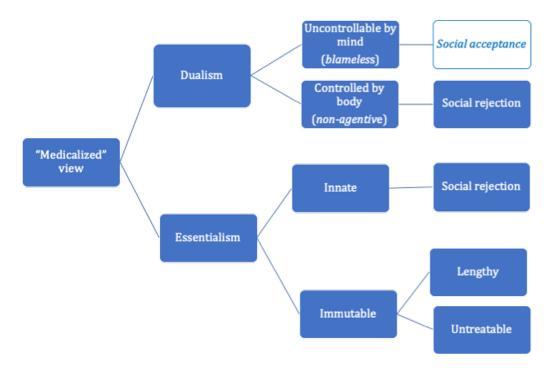


Figure 1. The role of Dualism and Essentialism in laypeople's attitudes towards "medicalized" psychiatric symptoms.

Essentialism is the intuitive belief that living things are what they are because they possess some immutable essence, and that this essence is transferred via biological inheritance from parents to offspring (Gelman, 2003; Gelman & Wellman, 1991; Keil, 1986; Medin & Ortony, 1989). For example, young children and adults believe that parents are more likely to share physical (Astuti, 2004; Gelman & Wellman, 1991; Hirschfeld, 1995; Solomon, Johnson, Zaitchik, & Carey, 1996) and psychological (e.g., Eidson & Coley, 2014; Heyman & Gelman, 2000) properties with their biological offspring than with adoptees who are raised in the same environment. Essentialism, then, guides our intuitive understanding of biological inheritance (Gelman & Wellman, 1991; Keil, 1986; Solomon et al., 1996). It suggests that certain traits form part of our innate immutable essence, and consequently, those traits are believed to be fixed (Gelman, 2004).

These properties of essentialist reasoning can explain laypeople's attitudes towards "medicalized" psychiatric conditions (Dar-Nimrod & Heine, 2011; Haslam & Ernst, 2002; Haslam & Kvaale, 2015). Indeed, if people view "medicalized" psychiatric conditions as innate, then, per Essentialism, laypeople would be further expected to consider these conditions as defining the patients' essence and thus immutable—and associated with symptoms that are lengthier (Lebowitz & Ahn, 2014; Lebowitz, Ahn, & Nolen-Hoeksema, 2013), less responsive to treatment (Kvaale et al., 2013b; Loughman & Haslam, 2018), and characteristic of patients' biological families (Bennett et al., 2008). And since Essentialism would suggest that the patient's core is different from one's own, Essentialism presents a second route for promoting social stigma (Kvaale et al., 2013a Loughman & Haslam, 2018).

To explain these negative public attitudes, all that is required, then, is for people to consider psychiatric conditions as innate; once they do, then, these essentialist projections should follow naturally. And indeed, many studies have explicitly informed participants that the psychiatric symptoms in question have genetic causes (Ahn, Bitran, & Lebowitz, 2020; Bennett et al., 2008; Boysen, 2011; Cheng, 2015; Lam, Salkovskis, & Warwick, 2005; Lebowitz & Ahn, 2014; Lebowitz et al., 2013; Walker & Read, 2002), so their perception as innate, immutable, and stigmatized is only expected.

Remarkably, similar negative attitudes and prognosis pessimism are observed even when disorders are strictly defined as "brain-based" (Loughman & Haslam, 2018). These results are notable because these experiments offer no evidence that the disorders in question are innate. Why, then, would such disorders trigger essentialist thinking? Do participants effectively presume that disorders that manifest in the brain reflect one's innate essence?

To be sure, such a presumption is false. Modern science tells us that *all* mental states—innate or learned—are brain states, so the detection of their correlates in the brain offers no evidence for innate origin. We suggest that this belief arises not from rational scientific analysis but from an oft-neglected aspect of intuitive Essentialism—the belief that the essence of living things is *embodied*.

#### b. The "embodied essence" hypothesis

Past research suggests that laypeople's perceptions of biological essence is linked to the body. For example, children state that a puppy is brown, like its mother, because it got a tiny *piece* of matter from its mother (Springer & Keil, 1991). Other results suggest that children believe that the essence of living things resides in their *insides* (Gelman & Wellman, 1991), that it must correspond to specific bodily *substance* (e.g., blood; Waxman, Medin, & Ross, 2007), and that it must be localized in a certain place—at their *center* (Newman & Keil, 2008). Discreteness (i.e., piece of matter), position in space (i.e., "insides"), and bodily substance are all properties that we intuitively project to biological matter, but not to ephemeral mental entities (e.g., to thoughts or ghosts). Together, these results open up the possibility that laypeople view the essence as part of the body. In other words, they believe that the innate essence of living things must be *embodied* (Berent, 2020a).

The hypothesis of an *embodied essence* is not new. This proposal goes back at least to Haslam and colleagues (2004), who hypothesized that people essentialize emotions because they perceive them as embodied. Similarly, Lindquist and colleagues (Lindquist et al., 2013) asserted that "categories whose instances are tied to the body (e.g., hunger) are more essentialized than are categories that are thought to exist in the mind (e.g., memory)" (Lindquist et al., 2013, p. 641). Nonetheless, this possibility has not been widely explored in the essentialist literature. We believe this hypothesis merits attention, as it has the potential to shed light on public attitudes towards mental disorders (additional applications are considered in the General Discussion).

If people believe that one's innate essence is embodied, then upon learning that a given trait is linked to the body, one would be more likely to essentialize that trait, hence, consider it as innate. Per the *embodiment hypothesis*, embodiment should thus spontaneously trigger essentialist thinking, even in the absence of evidence for a genetic cause.

If people are further Dualists, however, then they would not automatically conclude that every human trait is materially embodied. This is because the Dualist, recall, assumes that some traits (e.g., knowledge) are mental, ephemeral, and disembodied. So to consider a trait embodied (in line with Essentialism), the Dualist would require *explicit* evidence that such trait resides in the body. Brain tests provide the requisite proof. Behavioral evidence, in contrast, won't do, as behavioral outcomes could conceivably arise from the mind. Accordingly, upon learning that a given trait "shows up" in the brain, laypeople should be more likely to conclude that the trait in question is innate and immutable, courtesy of Dualism and Essentialism. As noted, this tendency is a bias, inasmuch as it is based on false premises, supplanted by Dualism and Essentialism, and it gives rise to conclusions that are at odds with science. As such, the presumption that *conditions that "show up" in the brain are innate* is an irrational psychological bias.

Our recent findings are in line with this hypothesis. First, when told that a given psychological trait can be detected in a brain scan, people are more likely to link that trait

with one's essence compared to when the same trait is detectable behaviorally (Berent & Platt, 2021). Second, when a given trait "shows up" in the brain, people are more likely to view it as innate (Berent, Barrett, & Platt, 2020; Berent, Platt, & Sandoboe, in press).

These results from typical psychological conditions, however, do not necessarily generalize to psychological disorders, nor do they speak to the question of immutability and stigma. Here we thus examine whether laypeople project similar misconceptions to psychiatric disorders.

Our investigation explores two questions. First, are people biased to presume that psychiatric disorders that are "in the brain" are innate—the key hallmark of one's innate essence (e.g., Gelman, 2003; Haslam et al., 2000; Keil, 1986)? Second, do people further consider such "brain disorders" as immutable and stigmatized?

#### c. The present study

To address these questions, we invited participants to reason about a patient's psychiatric condition based on the outcome of a psychological experiment whose results were gauged by either a behavioral test or a brain test. For example, to evaluate a patient for depression, the experiment compares her response to happy and sad faces. The two tests differed on how the patient's response is gauged. The behavioral test measured the speed of the patient's key response, whereas the brain test tracked her brain response (a characteristic spike). Participants were informed of the expected typical behavioral/brain response, and they were explicitly told that the patient's results suggested abnormality.

Critically, the brain and behavioral tests were strictly matched for their diagnostic value—all they suggested was whether or not the patient's response was abnormal; the brain test offered no additional information about brain localization or severity. Nonetheless, only the brain test offered explicit evidence that the disorder affects the brain. Of interest is whether people presume that disorders that patently manifest in the brain are more likely to be innate.

Experiment 1 evaluates the perceived innateness of brain disorders. Experiment 2 further examines whether brain disorders are presumed to be heritable and immutable, and explores social attitudes towards patients. If people essentialize brain disorders, then conditions that are diagnosed by the brain test should be considered more likely to be innate compared to those diagnosed by the behavioral test. Given that the "medicalized" view of psychiatric disorders can elicit both negative and positive social attitudes (see Figure 1), and that this medicalized approach can further remove blame (Haslam & Kvaale, 2015), it is not a priori clear whether the promotion of essentialist thinking by the brain test should further elicit stigma. However, we do expect conditions that are diagnosed in the brain to be considered as immutable—as lengthier and more difficult to treat.

## **Experiment 1**

#### Methods

Participants. Forty participants took part in Experiment 1. Participants were recruited from Amazon Mechanical Turk. They were all adult native English speakers who were reportedly free of language and reading disorders. Of all participants, 48% reported their highest completed level of education as high school, 45% as college, 8% as a graduate school program, and 0% reported completing none of the above education.

To be included in the sample, participants had to further provide a coherent explanation for their reasoning in the experiment; this requirement was adopted in order to eliminate bot responses. "Coherent" explanations were evaluated liberally: the explanation was acceptable as long as it offered some justification ("I went with my gut feelings") that was not copied verbatim from the vignette. To minimize the effect of prior education, we limited the sample of participants to those who had not taken advanced courses (i.e., beyond the Introductory level) in psychology or linguistics. Psychology and biology are of interest because these disciplines could directly shape participants' understanding of innateness and psychiatric disorders; the exclusion of linguistics students was imposed for extraneous reasons (the same IRB protocol included experiments related to language). 90% of participants indicated that they had not taken advanced courses in biology as well. Participants were paid \$0.80 for their participation, and the experiment lasted an average of nearly five minutes.

Sample size in Experiments 1-2 was informed by sensitivity power analysis of pilot results. These results suggested that the selected sample is sufficient to obtain a large effect size (.8) with a probability of .8 (and an alpha level of .05).

Materials and procedures. The materials consisted of four matched pairs of vignettes, loosely modeled after the materials in Lebowitz and Ahn (2014). Each such pair featured a female individual who sought treatment to alleviate some psychological symptoms; two pairs featured depression (e.g., Terry is a 28-year-old woman who is seeking treatment because she has felt deeply sad for the past 4 weeks); another two vignette pairs each featured a patient who suffered from social phobia (e.g., Michelle is a 21-year-old college student who has decided to seek treatment for what she calls "crippling shyness"). For each such pair (e.g., for the two depression patients), the vignettes were identical, except for the patient's identity.

Participants were told that a clinician suspects that the patient suffers from depression/social phobia and administers a psychological test in order to better understand the patient's condition—either a behavioral or brain test. Thus, within a single pair, each matched vignette presented either the behavioral or brain test.

Both tests compared the patient's responses to faces and inanimate objects by gauging either their behavioral response time or brain spikes (for the behavioral and brain

tests, respectively). Participants were informed of what outcomes are expected with typical individuals. For example, in the test for depression, the patient was presented with happy and sad faces; participants were told that typical individuals are expected to exhibit faster responses/brain spikes for happy faces compared to sad faces. Participants were told that individuals who suffer from a disorder (e.g., depression) are expected to show the opposite pattern (e.g., patients with depression should show faster responses/brain spikes to sad faces compared to happy faces). Participants were further informed that the patient's performance was in line with the pattern suggestive of a disorder (e.g., Terry showed a faster response/brain spike to sad faces relative to happy faces).

After reading each vignette, participants gauged the innateness of the disorders by rating (on a 1-7 scale) how likely they thought it would be that a close family member of the patient suffers from the same disorder (the specific relation to the patient—children, sister, or mother of the patient was varied across vignettes). All materials are provided in Appendix A.

Altogether, then, the materials included a total of 2 disorders (depression vs. social phobia) x 2 patients x 2 test (brain/behavior) vignettes. These eight vignettes were arranged in two counterbalanced lists, such that each such list included four vignettes (two vignettes for each disorder [with different patients], crossed with the two tests). Each participant was assigned to one of the two lists (for a total of 20 participants per list).

#### Results

Figure 2 presents the innateness means; in Experiments 1-2, error bars are 95% confidence intervals for the difference between the means; the scale's "neutral" midpoint is indicated by a dotted line.

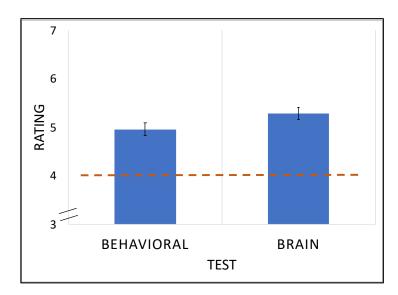


Figure 2. The perceived tendency of psychiatric disorders to run in patients' families (a gauge of innateness) given the brain and behavioral test results (in Experiment 1). The red dotted line represents the scale's neutral midpoint.

An inspection of the means suggests that people believed a disorder to be more likely to run in patients' families when the condition was diagnosed by a brain test compared to when precise the same disorder was diagnosed by a behavioral test.

These conclusions were confirmed by a 2 Test (brain/behavior) x 2 Disorder (depression/social phobia) fully repeated measures ANOVA. The main effect of Test was significant (F(1,39)= 6.39, p=.02,  $\eta^2_p$ =.14), as disorders diagnosed by the brain test were considered more likely to run in the family than disorders diagnosed by a behavioral test. There was also a reliable main effect of Disorder, as people considered social phobia (M=5.45) as more likely to run in the family than depression (M=4.80, F(1,39)= 15.66, p<.0004,  $\eta^2_p$ =.29). The interaction was not significant (F<1). Thus, people were more likely to attribute the symptoms to family members when the patient was diagnosed by a brain test.

## **Experiment 2**

The finding that disorders that manifest in the brain are perceived as more likely to run in patients' families (in Experiment 1) is in line with the hypothesis that brain disorders are viewed as defining one's innate essence. Families, however, can share psychological symptoms due to their shared environment, rather than shared genes. To secure the link between brain disorders and innateness, in Experiment 2 we asked people to evaluate whether the disorders in question would emerge in a biological family member who had no social contact with the patient. Of interest is whether the diagnosis of the condition by the brain test will increase its perceived innateness (i.e., heritability).

Experiment 2 further extended the results of Experiment 1 in three directions. First, we sought to determine whether the perception of innateness depends on the perception of the disorder as affecting the brain. To this end, we explicitly asked participants to indicate how likely they thought the condition was to affect the brain.

Second, we aimed to explore the link between the perceived innateness of these disorders with their perceived immutability and social attitudes towards patients. Immutability, here, was evaluated by the perceived duration of the symptoms and their susceptibility to treatment. To explore social attitudes, we further invited participants to evaluate their willingness to interact with the patient socially. One set of social measures examined close and consequential social interactions—as a marriage partner, as caretaker of one's child, and as a roommate; another set of measures examined interactions that are more remote and less personally consequential—as a coworker, as a friend, as a neighbor, or as a person to spend an evening with.

Third, we sought to demonstrate the generality of the results to additional psychiatric disorders. In principle, the disorder itself is inconsequential to our proposal, inasmuch as, per the *embodiment hypothesis*, brain tests should promote the perception of innateness across the board, irrespective of disorder. In practice, however, participants might be aware that certain psychiatric disorders run in families. Moreover, severe and

persistent psychiatric symptoms might trigger essentialist thinking by suggesting to people that the patient's essence is immutable and distinct from their own (Boysen, 2011; Haslam & Ernst, 2002; Kvaale et al., 2013a). Together, prior knowledge and severe symptoms could potentially counteract our test manipulation.

Given that the disorders studied in Experiment 1 are relatively mild and their heritability is modest (heritability estimates for depression: 22-37% (McGue & Christensen, 2003); for social phobia: 48% (Stein, Jang, & Livesley, 2002)), the question thus arises whether our previous findings could generalize even to conditions that are both severe and highly heritable. Schizophrenia and bipolar disorders present a case in point. Not only are these disorders highly heritable (heritability estimates for schizophrenia: 79% (Hilker et al., 2018); for bipolar disorder: 60-85% (Smoller & Finn, 2003), but a meta-analysis of the literature has shown that the effect of biogenetic explanations differs for schizophrenia and depression (Kvaale et al., 2013a). While for schizophrenia, biogenetic explanations were associated with the reduction of blame and an increase in the desire for social distancing, this was not the case for depression (Kvaale et al., 2013a). In Experiment 2, we thus invited participants to reason about schizophrenia and bipolar disorder; for comparison, we also included depression (as in Experiment 1)—a disorder that is less severe and heritable.

Given our pilot results (see SM), we expected that, when presented with such severe disorders, the effect test would be highly attenuated. To promote attention to the diagnostic test in Experiment 2, we presented each participant with two matched patients who suffered from the same symptoms—one was diagnosed by a brain test, another was diagnosed by a behavioral test (with order counterbalanced), and the narrative referenced this contrast repeatedly. In so doing, we sought to partly control for the inherent severity of the disorder and direct participants' attention to the test outcomes. For each such patient, participants were asked to evaluate whether the disorder likely affects the brain, to evaluate its innateness, its expected length, susceptibility to treatment, and social attitudes towards the patient.

We hypothesize that the diagnosis of such disorders by a brain test will promote essentialist thinking, in line with the *embodiment hypothesis*. We thus expect that the perception of psychiatric disorders as embodied in the brain will support their perception as innate and immutable, and decrease participants' willingness to engage in close social interactions with the patient.

#### Methods

Participants. Forty participants took part in Experiment 2. Participants were native English speakers, sampled from Prolific. Since Prolific does not support all the customized a priori filtering of participants used in Experiment 1, this sample was more heterogeneous than the MTurk sample in Experiment 1. Of all participants, 20% reported their highest completed level of education as high school, 45% as college, 35% as a graduate school program, and 0% reported completing none of the above education. Additionally,

participants reported taking coursework beyond an introductory course in psychology (83%), biology (40%), and linguistics (28%). Two participants reported having language/reading disorders. Of this sample, 73% of participants identified as female and 28% as male, and the average participant age was 25.95 years old. Participants were paid \$3.00 for their participation, and the experiment lasted an average of approximately 11 minutes.

*Materials and Procedure.* Participants read three vignettes, loosely modeled after the materials in Lebowitz & Ahn (2014) and Kim, Ahn, Johnson & Knobe (2016) (see Appendix B). Each vignette described the symptoms of one psychiatric disorder: either depression, schizophrenia, or bipolar disorder.

The vignette further introduced a pair of female patients (e.g., Terry and Jane) who suffered from the same set of symptoms. Participants were told that a clinician suspected a disorder (e.g., schizophrenia). To evaluate this diagnosis, the two patients were each presented with a standardized test. As in Experiment 1, the test featured a brief experiment that evaluated the patient's response. One patient in the pair was diagnosed by a brain test (a characteristic brain spike); the other was diagnosed by a matched behavioral test (the speed of their button pressing). Test order (brain/behavior) for each vignette was counterbalanced across two lists.

Participants were then informed of the results of each patient's test, and repeatedly reminded of its type (e.g., a behavioral test). They were explicitly told that the results for each patient suggested abnormality. Participants were then invited to respond to four sets of questions (some with sub-parts) concerning one of the patients (e.g., Terry). First, participants rated whether they believed the condition to affect the brain. Second, they rated the likelihood that the disorder would manifest in a close family member who was said to be related to the patient biologically, but had never met the patient (e.g., in an adoption situation); this is a gauge of the perceived innateness of the disorder. Third, participants rated the perceived duration of the symptoms, ease of treatment, and their susceptibility to treatment (using the questions from Cheng, 2015). Fourth, participants were asked to indicate their social attitudes towards the patient. Some of these questions probed for attitudes concerning *distant* social interactions, defined as participants' willingness to (a) have the patient move next door, (b) spend an evening with the patient, (c) make friends with them, (d) work with them closely. Others featured *close* social interactions, defined as participants' willingness to (e) have them marry into one's family, and (f) have them as a roommate; and (g) take care of one's child. Questions a-d were adopted from Link, Phelan, Bresnahan, Stueve, and Pescosolido (1999).

Next, participants considered the other member of the pair (e.g., Jane). Participants were reminded of the diagnostic test (e.g., brain test), and asked to respond to the four (multi-part) questions (as above).

As in Experiment 1, participants were encouraged to respond based on the test results alone. Responses to the "innateness," "affects brain," and "length" questions were given on a 1-7 scale (1=highly unlikely; 7=highly likely); for length (1=one week; 7=more

than three years); for "social attitudes," "ease of treatment," and "susceptibility to treatment" ratings were given on a 1-6 scale (as in Link et al., 1999; 1=definitely not, 6=definitely yes).

#### **Results and Discussion**

In what follows, we examine the effect of Test (brain vs. behavioral) on the perception of the disorder as innate, its effect on the brain, its immutability, and on social attitudes towards patients. Each such dependent measure is analyzed using a 2 Test (brain/behavior) x 3 Disorder (depression/schizophrenia/bipolar disorder) fully repeated measures ANOVA.

The analyses reported below focuses on effect of Test; Disorder is considered only inasmuch as it reliably modulates the contrast between matched tests for a single disorder (e.g., between the brain and behavioral tests for depression); all other effects of Disorder are described in the supplementary materials (SM). After analyzing the effect of Test, we next moved to consider the association between these various measures in a correlational analysis.

a. Innateness. Figure 3 plots the perceived innateness ratings. An inspection of the means suggests that people were more likely to consider the disorder as innate when the diagnosis was offered by the brain test relative to the behavioral test. In line with this conclusion, the ANOVA yielded a reliable main effect of Test (F(1,39)= 7.74, p=.01,  $\eta^2_p$ =.17), which was not further modulated by Disorder (F(2,78)= 2.43, p=.10,  $\eta^2_p$ =.06).

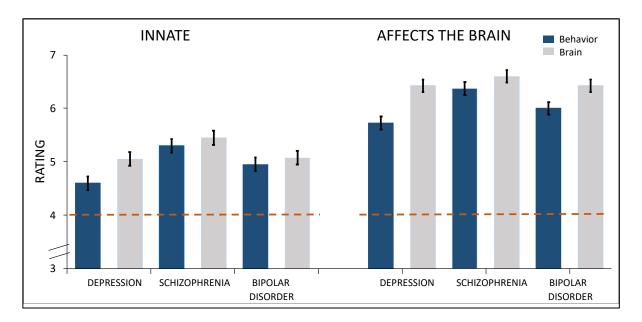


Figure 3. The perceived heritability of the disorder and its potential to affect the brain given the brain and behavioral tests (in Experiment 2). The red dotted line represents the scale's neutral midpoint.

b. Affects the brain. Figure 3 depicts the perceived potential of the disorder to affect the brain. The ANOVA yielded a reliable main effect of Test (F(1,39)= 7.92, p=.01,  $\eta^2_p$ =.17). As expected, people considered disorders diagnosed by a brain test as more likely to affect the brain.

The effect of Test, however, interacted with Disorder (F(2,78)= 3.59, p=.03,  $\eta^2_p$ =.08). Tukey HSD contrasts showed that, when the diagnosis was given by the brain test, people were more likely to perceive bipolar disorder (t(78)=2.51, p=.01) and depression (t(78)=4.03, p<0.0002) as affecting the brain compared to when these disorders were diagnosed behaviorally. This, however, was not the case for schizophrenia (t<1), possibly because participants tended to strongly associate this disorder with a brain disorder, even when the diagnosis was behavioral.

c. *Immutability*. Immutability was gauged by two measures: perceived length and susceptibility to treatment, assessed by the "ease of treatment" and "treatability" questions. Since the internal consistency between the two treatment questions was high (Cronbach alpha=.863), we conducted the analysis while collapsing across those two questions.

Considering length, an inspection of the means (see Figure 4) suggested that the brain test was associated with a lengthier disorder than the behavioral test. In line with this hypothesis, the ANOVA indeed yielded a reliable effect of Test (F(1,39)= 4.68, p=.04,  $\eta^2_p$ =.11), which was not further modulated by Disorder (F(2,78)= 223, p=.11,  $\eta^2_p$ =.05).

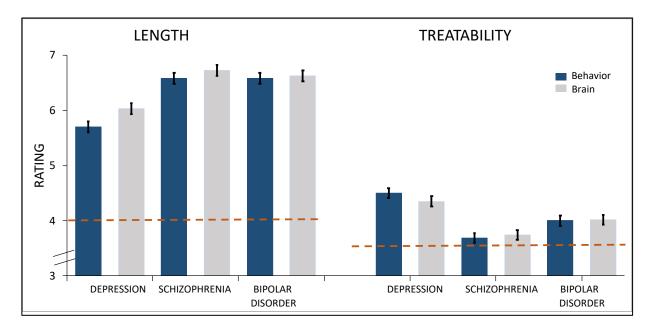


Figure 4. The perceived treatability and length of disorders diagnosed by brain vs. behavioral tests (in Experiment 2). The red dotted line represents the scale's neutral midpoint.

Similar analyses of the perceived treatability of the disorders (see Figure 3), however, did not find an effect of Test (F<1) or an interaction (F(2,78)= 1.13, p=.33,  $\eta^2_p$ =.03), possibly because participants did not consider these disorders as readily treatable.

d. Social attitudes. We next evaluated the effect of test on attitudes towards close vs. distant social interactions with patients. An inspection of the results (see Figure 5) suggests that people were not opposed to engaging in distant social interactions with patients (e.g., as a coworker), as the mean ratings were overall well above the scale's midpoint (3.5). But when it comes to closer social interactions (e.g., as a marriage partner), here participants were less willing to interact with the patient when the diagnosis was presented by a brain test. Additionally, their mean ratings were overall lower.

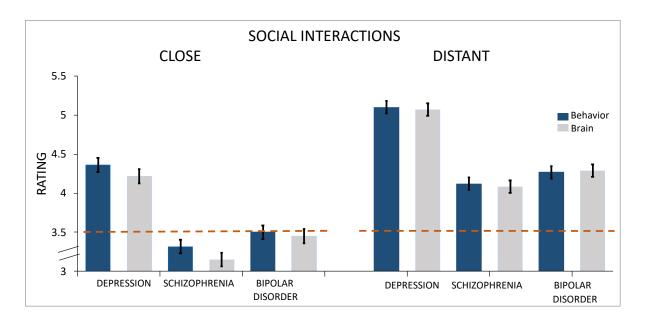


Figure 5. Willingness to engage in close and distant social interactions given the brain and behavioral tests in Experiment 2. The red dotted line represents the scale's neutral midpoint.

In line with this conclusion, the 2 Test x 3 Disorder ANOVA of attitudes towards distant social interaction did not yield a reliable effect of Test (F<1) nor was the Test x Disorder interaction significant (F<1).

In contrast, for close social interactions, the ANOVA yielded a reliable effect of Test  $(F(1,39)=5.23, p=.03, \eta^2_p=.12)$ , as participants were less willing to interact with patients diagnosed by the brain (relative to the behavioral) test. This effect was not further modulated by the effect of Disorder (F<1).

e. Correlational analysis. Our final set of analyses examined the association between the perceptions of the disorders as affecting the brain, as innate, as immutable, and participants' willingness to engage in social interactions with patients (either close or distant interactions) using a correlational analysis (averaged across the three disorders). Table 1 presents the results.

Results showed that the brain and behavioral tests elicited different correlational patterns. When the test was *behavioral*, disorders perceived to affect the brain were considered more likely to be innate and lengthier. Additionally, disorders that were perceived as innate were associated with lengthier progression. These same correlations were not significant when the diagnosis was given by the *brain* test, possibly because, here, all conditions were perceived more uniformly, as likely to affect the brain.

Table 1. The correlations between the perception of psychiatric conditions as affecting the brain and innate and stigma in Experiment 2, as indicated by the brain and behavioral tests

		Affects brain	Innate	Treatable	Lengthy	Social interactions(close)
Behavioral	Innate	0.46**				-
	Treatable	-0.13	0.04			
	Lengthy	0.70***	0.51***	0.05		
	Social interactions(close)	0.31	-0.16	0.34	0.09	
	Social interactions(distant)	0.20	-0.11	0.52	0.05	0.85
Brain	Innate	0.2				
	Treatable	0.13	0.21			
	Lengthy	0.2	0.02	0.24		
	Social interactions(close)	0.41	-0.22	0.38	0.15	
	Social interactions(distant)	0.25	-0.12	0.53	0.09	0.80

Note: \* = p < .05, \*\* = p < .01, \*\*\* = p < .001

Moving to social attitudes, here we found that, regardless of the diagnostic test, disorders perceived as treatable were associated with better social acceptance, and this was the case for both close and distant social interactions. Interestingly, when the disorder was perceived to affect the brain, participants were reportedly *more* willing to engage in close social interactions, and this was the case irrespective of the diagnostic test—contrary to the ANOVA result above, where a brain diagnosis yielded *lesser* willingness for close social interactions.

We speculate that these conflicting outcomes arose because a disorder can be perceived as "affecting the brain" for different reasons, and as noted (Figure 1), such "medicalized" accounts of the disorder can elicit conflicting social responses. One obvious clue for "brain status" is presented by the positive brain test, and as noted (see Figure 1), this clue could promote stigma by triggering essentialist thinking. Participants, however, could also arrive at the conclusion that a disorder affects the brain because they possess

prior knowledge suggesting that the disorders in question have a scientific brain explanation (irrespective of the diagnostic test presented to them). Because this conclusion does not arise from Essentialism, it is not expected to trigger stigma, and since it removes blame from patients, this perspective could *improve* social attitudes (see also Haslam & Kvaale, 2015).

Summarizing, Experiment 2 found that when participants were provided with evidence that the disorder could be detected in the brain (using the brain test), they were more likely to consider it as innate and lengthier compared to when the *same condition* was diagnosed behaviorally, and correspondingly, they were less willing to closely interact with the patient (as a marriage partner, as a caretaker for one's child, and as a roommate). Moreover, the correlational analysis found that when participants spontaneously perceived the disorders as embodied in the brain (as the test was behavioral), they tended to consider them as lengthier and innate. These results are all in line with the hypothesis that the brain diagnosis promotes essentialist thinking.

These conclusions, however, are subject to several limitations. First, for the severe disorders in Experiment 2, we only observed the effect of test when participants were prompted to explicitly compare the two tests to each other (by contrasting among matched patients, diagnosed by distinct tests). This observation is further bolstered by the results of Experiments 3-5, presented in the SM. These experiments demonstrate that the effect of test varied systematically, depending on the severity of the disorder and the experimental design. For milder disorders, such as depression and social phobia (in Experiment 5, see SM), brain tests increased the perception of the disorder as innate, and this was the case even when participants did not directly contrast between the two tests (when participants simply considered each patient and her test result individually, using the same design as in Experiment 1). Severe disorders (schizophrenia, bipolar disorder), however, were inherently perceived to affect the brain, and this was the case even when the test results were behavioral. So for these severe disorders, the brain test did not prompt the perception of the disorder as innate unless participants explicitly contrasted the brain and behavioral tests (as in Experiment 2). Less direct manipulations did not elicit an effect of test, and this was the case regardless of whether test was manipulated between- or within participants (in Experiments 3-4). This is one notable limitation of the present conclusions.

Additionally, some of the predictions of essentialist reasoning were not borne out by our results. First, the brain test did not affect the perception of treatability. Second, treatability did not correlate with the perceived effect on the brain, nor did it correlate with length. These outcomes are at odds with the notion that treatability and length are both gauges of immutability. We speculate that these null outcomes could have emerged because participants did not perceive these severe disorders as highly treatable, but this possibility ought to be independently evaluated. Another challenge to the essentialist account is presented by the finding that disorders perceived to affect the brain were associated with a greater willingness to engage in close social interactions with patients, possibly because these perceptions were based on prior information about the disorder (rather than by essentialist reasoning).

Notwithstanding these challenges, the results of Experiment 2 converge with those of Experiment 1 to suggest that disorders that are linked to the brain are perceived as immutable and lengthier.

#### **General Discussion**

A large public campaign has sought to destignatize psychiatric disorders by promoting the view that they are brain disorders. But these efforts have not eradicated the stigma. In fact, psychiatric disorders with neurobiological correlates are associated with more negative attitudes and poorer perceived prognoses compared to psychosocial symptoms (Loughman & Haslam, 2018). The present research has examined whether these puzzling attitudes could be due to the essentialist presumption that brain disorders are innate.

Experiments 1-2 have shown that laypeople are more likely to view psychiatric disorders as innate when the diagnosis is supported by a brain test compared to a matched behavioral test. In Experiment 2, participants further considered disorders diagnosed by a brain test as lengthier, and as expected, they also were more likely to classify these conditions as "affecting the brain." A correlational analysis indicated that, when the brain status of the disorder was ambiguous (i.e., in the behavioral test), conditions perceived to affect the brain were associated with innate origin and lengthier course—a result suggesting immutability.

These results, however, are subject to several limitations. First, the perceived effect of the disorder on the brain did not affect the perception of the condition as treatable (another gauge of immutability), nor did length and treatability correlate, possibly because these disorders were overall considered severe and untreatable. Second, for severe heritable disorders, like schizophrenia and bipolar disorders (in Experiment 2), sensitivity to the test only obtained when participants were invited to explicitly compare two matched patients, diagnosed by distinct tests; less severe disorders did not require this explicit contrast (in Experiment 1; see also SM). Nonetheless, our results are mostly in line with the possibility that conditions associated with the brain are considered as innate and immutable.

Moving to the social consequences of the "medicalized" view, here our results were mixed. On the one hand, when the disorder was diagnosed by the brain test, people were less willing to interact closely with the patient (e.g., as a marriage partner). On the other hand, we found that the perception of the condition as "affecting the brain" was associated with a greater willingness to interact with the patient closely. We will return to this mixture of social attitudes below.

Altogether, then, our results suggest that when a psychiatric condition is diagnosed by the brain, participants consider it as innate, they associate it with a longer progression, and with a mixture of social attitudes, including evidence for stigma.

Innateness, immutability, and stigma are all known attributes of Essentialism. Accordingly, a large literature has linked these reactions to Essentialism (for reviews and meta-analyses: Ahn et al., 2017; Dar-Nimrod & Heine, 2011; Kvaale et al., 2013a; Kvaale et al., 2013b; Loughman & Haslam, 2018). As noted, however, many previous studies have explicitly informed participants of a genetic origin, and this information could explain the essentialist bias. The novel finding here is that participants project these same attitudes even when the evidence presented to them offered absolutely no information about the innate origin of the disorder. Thus, our results show for the first time that people spontaneously presume that if psychiatric disorders are linked to the brain, those conditions are immutable and innate.

The tendency to view conditions diagnosed by a brain test as innate cannot be explained by the greater informative value of the brain test. As noted, the brain and behavioral tests offered precisely the same diagnostic information, and they were paired with the same symptoms.

It is also unlikely that participants were more likely to accept the brain results as innate because they believe brain-based disorders to be better understood by the scientific community, or from a desire to enforce congruence between the "scientific" test result and a "scientific" level of analysis. This explanation assumes that people favor non-reductive explanations—ones that situate the explanation and the datum within the same level of analysis. But an investigation of laypeople's explanatory preferences suggests just the opposite: people normally prefer reductive explanations to non-reductive ones, and this preference is strongest when behavior is reduced to a brain explanation (Hopkins, Weisberg, & Taylor, 2016).

Could our findings, then, arise because laypeople consider brain explanations to be more satisfying than behavioral explanations (Hopkins et al., 2016; Weisberg et al., 2008; Weisberg, Taylor, & Hopkins, 2015)? We believe not. Indeed, an improved understanding (either real or presumed) does not imply that the disorders are innate and immutable, nor can it account for the correlation between the perceived effect of the disorder on the brain and its innateness and length.

Why, then, do people fall for this irrational bias? As noted, misconceptions about psychiatric disorders could arise from intuitive psychology, from Dualism and Essentialism (for reviews: Ahn et al., 2017; Dar-Nimrod & Heine, 2011; Haslam & Kvaale, 2015; Loughman & Haslam, 2018). Essentialism, specifically, explains why psychiatric disorders are viewed as immutable (Ahn et al., 2017; Kvaale et al., 2013b; Loughman & Haslam, 2018), on the one hand, and as likely to run in the patient's family (Bennett et al., 2008), on the other. As such, existing essentialist theories outline the mechanism that promotes the perception of essentialized traits as innate and immutable. The open question, though, is what triggers this essentialist thinking: why are people more likely to essentialize conditions that manifest in the brain?

To explain this puzzle, we hypothesize that people presume that the innate essence of living things is materially *embodied* (Berent, 2020a; Berent, 2020b). If people make this

presumption, then they would conclude that disorders that are embodied are potentially indicative of a person's essence. Behavioral tests, however, do not offer evidence for embodiment. This is because laypeople are not only Essentialists; they are also Dualists (Bloom, 2004)—they view the mind as immaterial, distinct from the body. Accordingly, upon learning that a given condition affects behavior, people do not automatically conclude that the trait in question is embodied. By contrast, brain explanations present explicit evidence for embodiment. So, if per Essentialism, innate traits must be materially embodied, then, upon learning that the condition "shows up" in the brain, the Dualist is now offered evidence that the trait in question potentially defines one's innate immutable essence.

The presumption of an embodied essence and Dualism can further account for the mixed social attitudes we have observed. As noted, by attributing the patient's actions to their body rather than their mind (their psychological core), per Dualism, the "medicalized" view can remove blame from the patient, hence, improve acceptance. In so doing, however, Dualism is further expected to challenge the human agency of the patient, and thus, promote negative attitudes. Essentialism could also exacerbate these negative attitudes by suggesting that the patient's essence is distinct from one's own. Together, Dualism and Essentialism are expected to elicit conflicting social attitudes (see Figure 1). As noted, however, the reduction in blame could also arise for reasons unrelated to intuitive reasoning—because participants are aware of scientific discussions of the disorder, and believe that it is uncontrollable by the patient (Haslam & Kvaale, 2015). Either way, social attitudes are expected to be mixed, and this complexity is indeed borne out by our findings. The embodied essence hypothesis can help explain these findings.

The embodied essence hypothesis is not new (for related proposals, see Haslam et al., 2004; Lindquist et al., 2013; Newman & Keil, 2008). Previous developmental research has shown that reasoning about biological inheritance anchors the essence of living things in the body (Gelman & Wellman, 1991; Newman & Keil, 2008; Springer & Keil, 1991). Recent results found the same in reasoning about psychological traits. Specifically, people are more likely to essentialize (typical) psychological traits—either emotions (Berent et al., 2020) or epistemic and sensorimotor traits (Berent et al., in press)—when these traits are diagnosed by a brain test compared to a matched behavioral test. Furthermore, the stronger the perceived anchoring of a psychological trait in the body, the more likely it is to be perceived as innate (Berent et al., 2020; Berent et al., in press). The present results suggest that the presumption of an embodied essence could further taint our understanding of psychiatric disorders.

The links we have uncovered between brain explanations and perceptions of the innateness of disorders, their length, and social attitudes towards patients can also explain why neuroscience has not eradicated public misconceptions about psychiatric patients (Ahn et al., 2017; Bennett et al., 2008; Haslam & Kvaale, 2015; Kvaale et al., 2013a; Kvaale et al., 2013b; Lebowitz & Ahn, 2014; Loughman & Haslam, 2018). In fact, these results open up the possibility that the medicalized view of psychiatric disorders could have partly backfired.

At a yet broader level, Essentialism could further explain some of the allure of brain science for the public (e.g., Weisberg et al., 2008). If we identify brain responses with our innate essence, then it is no wonder we assign brain findings special significance and disregard their logical inconsistencies. Such intuitive cognitive biases can derail scientific reasoning (Gottlieb & Lombrozo, 2018; Shtulman, 2017) and obscure our ability to grasp the workings of our own minds (Berent, 2020a).

## References

- Ahn, W.-k., Bitran, A., & Lebowitz, M. (2020). Effects of genetic information on memory for severity of depressive symptoms. *Plos One, 15*(10), e0239714-e0239714. doi:10.1371/journal.pone.0239714
- Ahn, W.-k., Kim, N., S., & Lebowitz, M., S. (2017). The Role of Causal Knowledge in Reasoning About Mental Disorders. In M. Waldmann, R. (Ed.), Oxford library of psychology. The Oxford handbook of causal reasoning (pp. 603–617): Oxford University Press.
- Astuti, R. (2004). Constraints on conceptual development: a case study of the acquisition of folkbiological and folksociological knowledge in Madagascar. Boston, MA: Boston, MA: Blackwell Pub.
- Bennett, L., Thirlaway, K., & Murray, A. J. (2008). The Stigmatising Implications of Presenting Schizophrenia as a Genetic Disease. *Journal of Genetic Counseling*, *17*(6), 550-559. doi:10.1007/s10897-008-9178-8
- Berent, I. (2020a). *The blind storyteller: how we reason about human nature*: Oxford University Press.
- Berent, I. (2020b). On the matter of essence. Manuscript submitted for publication.
- Berent, I., Barrett, L. F., & Platt, M. (2020). Essentialist biases in reasoning about emotions. Frontiers In Psychology: Cognitive Science(23). doi:https://doi.org/10.3389/fpsyg.2020.562666
- Berent, I., & Platt, M. (2021). The true "me"—mind or body? *Journal of Experimental Social Psychology*, 93.
- Berent, I., Platt, M., & Sandoboe, G. M. (in press). Empiricism is natural: it arises from dualism and essentialism. *Oxford Studies in Experimental Philosophy*.
- Bloom, P. (2004). *Descartes' baby: how the science of child development explains what makes us human.* New York: Basic Books.
- Boysen, G. A. (2011). Biological Explanations and Stigmatizing Attitudes: Using Essentialism and Perceived Dangerousness to Predict Antistigma Intervention Effectiveness. *J Soc Psychol*, 151(3), 274-291. doi:10.1080/00224545.2010.481689
- Cheng, Z. H. (2015). Asian Americans and European Americans' stigma levels in response to biological and social explanations of depression. *Social Psychiatry and Psychiatric Epidemiology*, *50*(5), 767-776. doi:10.1007/s00127-014-0999-5
- Dar-Nimrod, I., & Heine, S. J. (2011). Genetic Essentialism: On the Deceptive Determinism of DNA. *Psychological Bulletin*, *137*(5), 800-818. doi:10.1037/a0021860
- Deacon, B. J., & Baird, G. L. (2009). The chemical imbalance explanation of depression: Reducing blame at what cost? *Journal of Social and Clinical Psychology*, 28(4), 415-435. doi:10.1521/jscp.2009.28.4.415
- Eidson, R. C., & Coley, J. D. (2014). Not so fast: Reassessing gender essentialism in young adults. *Journal of Cognition and Development, 15*(2), 382-392. doi:10.1080/15248372.2013.763810
- Gelman, S. A. (2003). *The essential child : origins of essentialism in everyday thought*. Oxford; New York: Oxford University Press.

- Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*(9), 404-409. doi:http://dx.doi.org/10.1016/j.tics.2004.07.001
- Gelman, S. A., & Wellman, H. M. (1991). Insides and essence: Early understandings of the non-obvious. *Cognition*, 38(3), 213-244. doi:10.1016/0010-0277(91)90007-Q
- Gottlieb, S., & Lombrozo, T. (2018). Can Science Explain the Human Mind? Intuitive Judgments About the Limits of Science. *Psychological Science*, *29*(1), 121-130. doi:10.1177/0956797617722609
- Greene, J., & Cohen, J. (2004). For the Law, Neuroscience Changes Nothing and Everything. *Philosophical Transactions: Biological Sciences, 359*(1451), 1775-1785. doi:10.1098/rstb.2004.1546
- Haslam, N. (2006). Dehumanization: An Integrative Review. *Personality and Social Psychology Review*, 10(3), 252-264. doi:10.1207/s15327957pspr1003 4
- Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist Beliefs About Personality and Their Implications. *Personality and Social Psychology Bulletin, 30*(12), 1661-1673. doi:10.1177/0146167204271182
- Haslam, N., & Ernst, D. (2002). Essentialist beliefs about mental disorders. *Journal of Social and Clinical Psychology*, 21(6), 628-644. doi:10.1521/jscp.21.6.628.22793
- Haslam, N., & Kvaale, E. P. (2015). Biogenetic Explanations of Mental Disorder: The Mixed-Blessings Model. *Current Directions in Psychological Science*, *24*(5), 399-404. doi:10.1177/0963721415588082
- Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*(1), 113-127. doi:10.1348/014466600164363
- Heyman, G. D., & Gelman, S. A. (2000). Beliefs About the Origins of Human Psychological Traits. *Developmental Psychology*, *36*(5), 663-678. doi:10.1037/0012-1649.36.5.663
- Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge, T. M., . . . Glenthøj, B. (2018). Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biol Psychiatry*, 83(6), 492-498. doi:10.1016/j.biopsych.2017.08.017
- Hirschfeld, L. A. (1995). Do children have a theory of race? Cognition, 54(2), 209-252.
- Hopkins, E. J., Weisberg, D. S., & Taylor, J. C. V. (2016). The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition*, 155, 67-76. doi:10.1016/j.cognition.2016.06.011
- Keil, F. C. (1986). The acquisition of natural kind and artifact term. In W. Demopoulos & A. Marras (Eds.), *Language Learning and Concept Acquisition* (pp. 133-153). New Jersey: Ablex: Norwood,.
- 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. doi:10.1016/j.brat.2014.02.009
- Kim, N. S., Ahn, W.-K., Johnson, S. G. B., & Knobe, J. (2016). The Influence of Framing on Clinicians' Judgments of the Biological Basis of Behaviors. *Journal of Experimental Psychology: Applied*, 22(1), 39-47. doi:10.1037/xap0000070
- Kvaale, E. P., Gottdiener, W. H., & Haslam, N. (2013a). Biogenetic explanations and stigma: A meta-analytic review of associations among laypeople. *Social Science & Medicine*, *96*, 95-103. doi:https://doi.org/10.1016/j.socscimed.2013.07.017

- Kvaale, E. P., Haslam, N., & Gottdiener, W. H. (2013b). The 'side effects' of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review*, *33*(6), 782-794. doi:https://doi.org/10.1016/j.cpr.2013.06.002
- Lam, D. C. K., Salkovskis, P. M., & Warwick, H. M. C. (2005). An experimental investigation of the impact of biological versus psychological explanations of the cause of "mental illness". *Journal of mental health (Abingdon, England)*, 14(5), 453-464. doi:10.1080/09638230500270842
- 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. doi:10.1073/pnas.1414058111
- 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), 518-527. doi:10.1037/a0031730
- Lindquist, K. A., Gendron, M., Oosterwijk, S., & Barrett, L. F. (2013). Do people essentialize emotions? Individual differences in emotion essentialism and emotional experience. *Emotion*, *13*(4), 629-644. doi:10.1037/a0032283
- 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. *Am J Public Health*, 89(9), 1328-1333. doi:10.2105/ajph.89.9.1328
- Loughman, A., & Haslam, N. (2018). Neuroscientific explanations and the stigma of mental disorder: a meta-analytic study. (Original Article) (Report). *Cognitive Research: Principles and Implications*, 3(1). doi:10.1186/s41235-018-0136-1
- McGue, M., & Christensen, K. (2003). The Heritability of Depression Symptoms in Elderly Danish Twins: Occasion-Specific Versus General Effects. *Behav Genet, 33*(2), 83-93. doi:10.1023/A:1022545600034
- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), Similarity and analogical reasoning. (pp. 179-195). New York, NY: Cambridge University Press.
- Newman, G. E., & Keil, F. C. (2008). Where is the essence? Developmental shifts in children's beliefs about internal features. *Child Development*, *79*(5), 1344-1356. doi:10.1111/j.1467-8624.2008.01192.x
- Nichols, S. (2011). Experimental philosophy and the problem of free will. *Science*, *331*(6023), 1401. doi:10.1126/science.1192931
- Pescosolido, B. A., Martin, J. K., Long, J. S., Medina, T. R., Phelan, J. C., & Link, B. G. (2010). "A disease like any other"? A decade of change in public reactions to schizophrenia, depression, and alcohol dependence.(Author abstract)(Report). *American Journal of Psychiatry*, 167(11), 1321. doi:10.1176/appi.ajp.2010.09121743
- Schomerus, G., Schwahn, C., Holzinger, A., Corrigan, P. W., Grabe, H. J., Carta, M. G., & Angermeyer, M. C. (2012). Evolution of public attitudes about mental illness: a systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*, *125*(6), 440-452. doi:10.1111/j.1600-0447.2012.01826.x
- Shtulman, A. (2017). Scienceblind: why our intuitive theories about the world are so often wrong: New York: Basic Books.

- Smoller, J. W., & Finn, C. T. (2003). Family, twin, and adoption studies of bipolar disorder. *Am J Med Genet C Semin Med Genet*, 123C(1), 48-58. doi:10.1002/ajmg.c.20013
- Solomon, G. E., Johnson, S. C., Zaitchik, D., & Carey, S. (1996). Like father, like son: young children's understanding of how and why offspring resemble their parents. *Child Development*, *67*(1), 151-171.
- Springer, K., & Keil, F. C. (1991). Early Differentiation of Causal Mechanisms Appropriate to Biological and Nonbiological Kinds. *Child Development*, 62(4), 767. doi:10.2307/1131176
- Stein, M. B., Jang, K. L., & Livesley, W. J. (2002). HERITABILITY OF SOCIAL ANXIETY-RELATED CONCERNS AND PERSONALITY CHARACTERISTICS: A TWIN STUDY. *J Nerv Ment Dis*, 190(4), 219-224. doi:10.1097/00005053-200204000-00002
- Strohminger, N., Knobe, J., & Newman, G. (2017). The True Self: A Psychological Concept Distinct From the Self. *Perspectives on Psychological Science*, *12*(4), 551-560. doi:10.1177/1745691616689495
- United States. Public Health Service. Office of the Surgeon, G., Center for Mental Health, S., & National Institute of Mental, H. (1999). Mental health: a report of the Surgeon General. Rockville, Md.: Pittsburgh, PA: Rockville, Md.: Dept. of Health and Human Services, U. S. Public Health Service.
- Walker, I., & Read, J. (2002). The differential effectiveness of psychosocial and biogenetic causal explanations in reducing negative attitudes toward "mental illness". *Psychiatry*, 65(4), 313-325.
- Waxman, S., Medin, D., & Ross, N. (2007). Folkbiological reasoning from a cross-cultural developmental perspective: early essentialist notions are shaped by cultural beliefs. *Developmental Psychology*, *43*(2), 294-308.
- Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The Seductive Allure of Neuroscience Explanations. *Journal of Cognitive Neuroscience*, *20*(3), 470-477. doi:10.1162/jocn.2008.20040
- Weisberg, D. S., Taylor, J. C. V., & Hopkins, E. J. (2015). Deconstructing the seductive allure of neuroscience explanations. *Judgment & Decision Making*, 10(5), 429-441.

## Appendix A

## Materials for Experiment 1

#### **General Introduction**

Psychological disorders can arise from various distinct sources—some conditions arise primarily with one's life experiences, whereas others result from genetic mutations that can affect other family members.

In this experiment, we describe several patients who underwent psychological tests in order to diagnose their conditions. The test results suggested that the person indeed exhibits an abnormality.

Given these test results, here, we ask you to reason about how likely it is that the disorder affects the patient's close relatives. We ask that you please make this determination based *only* on the test results.

#### **Depression 1: Behavioral test**

Terry is a 28-year-old woman who has felt deeply sad for the past 4 weeks, so she has decided to undergo a clinical evaluation.

A clinician suspects Terry might be suffering from depression. To better understand her condition, he administers Terry a psychological test. The test presents her with various types of images. Some images depict a human face--either neutral, happy, or sad; other images depict inanimate objects (e.g., an apple). Terry is simply asked to press a button to indicate whether or not the visual stimulus is a face.

Past research has shown that, in individuals without depression, happy faces elicit faster responses, whereas people with depression show the opposite pattern (faster responses to sad faces). Terry responds faster to the sad faces, suggesting that her condition is abnormal.

Terry is now considering getting pregnant, but she is worried that her children might exhibit similar symptoms.

Given Terry's test results, we ask you to determine how likely it is that Terry's future children will exhibit similar symptoms. Please base your answer *only* on the results of the test.

1=highly unlikely; 2=unlikely; 3=somewhat unlikely; 4=neither likely nor unlikely; 5=somewhat likely; 6=likely; 7=highly likely

#### **Depression 1: Brain test**

Terry is a 28-year-old woman who has felt deeply sad for the past 4 weeks, so she has decided to undergo a clinical evaluation.

A clinician suspects Terry might be suffering from depression. To better understand her condition, he administers Terry a psychological test. The test presents her with various types of images. Some images depict a human face--either neutral, happy or sad; other images depict inanimate objects (e.g., an apple). As Terry watches the images, the activation of her brain is recorded.

Past research has shown that, in individuals without depression, happy faces elicit a spike in brain responses, whereas people with depression show the opposite pattern (a brain spike to sad faces). Terry shows a spike in brain responses to the sad faces, suggesting that her condition is abnormal.

Terry is now considering getting pregnant, but she is worried that her children might exhibit similar symptoms.

Given Terry's test results, we ask you to determine how likely it is that Terry's future children will exhibit similar symptoms. Please base your answer *only* on the results of the test.

#### **Depression 2: Behavioral test**

Alex is a 39-year-old woman who is seeking treatment because she has reportedly "lost the ability to take pleasure in anything".

A clinician suspects Alex might be suffering from depression. To better understand her condition, he administers Alex a psychological test. The test presents her with various types of images. Some images depict a human face--either neutral, happy or sad; other images depict inanimate objects (e.g., an apple). Alex is simply asked to press a button to indicate whether or not the visual stimulus is a face.

Past research has shown that in individuals without depression, happy faces elicit a faster response, whereas people with depression show the opposite pattern (faster response to sad faces). Alex responds faster to the sad faces, suggesting that her test results are abnormal.

Alex has a teenager daughter who has recently started getting withdrawn and sad; she has been experiencing these symptoms for several months.

Given Alex's test results, how likely it is that her daughter suffers from the same condition suffers from the same condition. Please base your answer *only* on the results of the test.

1=highly unlikely; 2=unlikely; 3=somewhat unlikely; 4=neither likely nor unlikely; 5=somewhat likely; 6=likely; 7=highly likely

#### **Depression 2: Brain test**

Alex is a 39-year-old woman who is seeking treatment because she has reportedly "lost the ability to take pleasure in anything".

A clinician suspects Alex might be suffering from depression. To better understand her condition, he administers Alex a psychological test. The test presents her with various types of images. Some images depict a human face--either neutral, happy or sad; other images depict inanimate objects (e.g., an apple). As Alex watches the images, the activation of her brain is recorded.

Past research has shown that, in individuals without depression, happy faces elicit a spike in brain responses, whereas people with depression show the opposite pattern (a brain spike to sad faces). Alex shows a spike in brain responses to the sad faces, suggesting that her test results are abnormal.

Alex has a teenager daughter who has recently started getting withdrawn and sad; she has been experiencing these symptoms for several months.

Given Alex's test results, how likely it is that her daughter suffers from the same condition suffers from the same condition. Please base your answer *only* on the results of the test.

#### Social Phobia 1: Behavioral test

Michelle is a 21-year-old college student who experiences what she calls "crippling shyness"; she thus decides to get tested in order to diagnose her condition.

A clinician suspects Michelle might be suffering from a social phobia. To better understand her condition, he administers Michelle a psychological test. The test presents her with various types of images—either human faces or objects. Michelle is simply asked to press a button to indicate whether or not the visual stimulus is a face.

Past research has shown that, in individuals without social phobia, faces elicit faster responses, whereas people with social phobia show the opposite pattern (faster responses to objects). Michele responds faster to objects, suggesting that her results are abnormal.

Michelle shares the results with her sister Jane, who suffers from similar symptoms. Given Michelle's test results, we ask you to determine how likely it is that her sister suffers from the same condition. Please base your answer *only* on the results of the test.

1=highly unlikely; 2=unlikely; 3=somewhat unlikely; 4=neither likely nor unlikely; 5=somewhat likely; 6=likely; 7=highly likely

#### Social Phobia 1: Brain test

Michelle is a 21-year-old college student who experiences what she calls "crippling shyness"; she thus decides to get tested in order to diagnose her condition.

A clinician suspects Michelle might be suffering from a social phobia. To better understand her condition, he administers Michelle a psychological test. The test presents her with various types of images --either human faces or objects. As Michelle watches these images, the activation of her brain is recorded.

Past research has shown that, in individuals without social phobia, faces elicit a spike in brain responses, whereas people with social phobia show the opposite pattern (a brain spike to objects). Michele shows a spike in brain responses to objects, suggesting that her results are abnormal.

Michelle shares the results with her sister Jane, who suffers from similar symptoms. Given Michelle's test results, we ask you to determine how likely it is that her sister suffers from the same condition. Please base your answer *only* on the results of the test.

#### Social phobia 2: Behavioral test

Nicole is a 22-year-old recent college graduate who is seeking treatment because she becomes anxious whenever she must interact with strangers.

A clinician suspects Nicole might be suffering from a social phobia. To better understand her condition, he administers Nicole a psychological test. The test presents her with various types of images—either human faces or objects. Nicole is simply asked to press a button to indicate whether or not the visual stimulus is a face.

Past research has shown that, in individuals without social phobia, faces elicit a faster response, whereas people with social phobia show the opposite pattern (faster responses to objects). Nicole responds faster to objects, suggesting that her results are abnormal.

Nicole talks to her mother on the phone and tells her about the test's outcomes. As it turns out, Nicole's mother is likewise unusually anxious interacting with strangers. Given Nicole's test results, how likely it is that her mother suffers from the same condition. Please base your answer *only* on the results of the test.

1=highly unlikely; 2=unlikely; 3=somewhat unlikely; 4=neither likely nor unlikely; 5=somewhat likely; 6=likely; 7=highly likely

#### Social phobia 2: Brain test

Nicole is a 22-year-old recent college graduate who is seeking treatment because she becomes anxious whenever she must interact with strangers.

A clinician suspects Nicole might be suffering from a social phobia. To better understand her condition, he administers Nicole a psychological test. The test presents her with various types of images -- either human faces or objects. As Nicole watches these images, the activation of her brain is recorded.

Past research has shown that, in individuals without social phobia, faces elicit a spike in brain responses, whereas people with social phobia show the opposite pattern (a brain spike to objects). Nicole shows a spike in brain responses to objects, suggesting that her results are abnormal.

Nicole talks to her mother on the phone and tells her about the test's outcomes. As it turns out, Nicole's mother is likewise unusually anxious interacting with strangers. Given Nicole's test results, how likely it is that her mother suffers from the same condition. Please base your answer *only* on the results of the test.

## Appendix B

## Materials for Experiment 2

#### **General Introduction**

In this experiment, we are asking you to advise a clinician who is presented with three pairs of patients. The clinician suspects that each patient suffers from a condition, and he administers various tests to evaluate the patient's symptoms. We ask you to help the clinician judge the patient's condition.

#### **Depression vignette:**

Terry and Jane are two 28-year-old women who are seeking treatment because they each have felt deeply sad for the past 4 weeks. Terry and Jane state that ever since the current period of sadness began, they have had difficulty with memory and concentration.

A clinician suspects the two women might be suffering from depression. To better understand their condition, he administers each woman a psychological test.

The test presents the patient with various types of images. Some images depict a human face--either neutral, happy, or sad; other images depict inanimate objects (e.g., an apple). The test examines the patient's response to these images. However, the tests administered to Terry and Jane differ on how their response is measured.

**Terry's** response is evaluated by asking her to **press a button** to indicate whether or not the visual stimulus is a face. **Jane's response** is evaluated by recording the activation of her **brain**.

Both tests are strictly normed, and the clinician knows that for people with depression, sad faces elicit either faster button-pressing responses or brain spikes relative to happier faces (unlike people without depression).

The clinician has determined that Terry and Jane's responses to the test each exhibit the abnormal pattern. But as noted for Terry, this conclusion is based on the result of her behavioral (button-pressing) response, whereas for Jane, the conclusion is based on her brain response (brain spike).

With this information in mind, consider these various questions about Terry (diagnosed by the **behavioral** test):

- 1. Some psychiatric disorders result in clear changes in the brain, whereas others, the effects are evident psychologically. How likely is it that Terry's condition affects her brain?
- 2. Suppose now that Terry is pregnant. Because she does not have a partner, she wants to give her future daughter up for adoption. Terry would have no contact with her. Given Terry's test results, how likely do you think it is that, as the daughter matures, she will exhibit the same condition as Terry?
- 3. Please now consider the following:
  - a. How treatable do you think Terry's condition will be?
  - b. How easy will it be to treat Terry's condition?
  - c. How long do you think Terry will have this condition (1 week, 1 week to 1 month, 1 month to 3 months, 3 months to 6 months, 6 months to 12 months, 1 year to 3 years, more than 3 years
- 4. Suppose you met Terry. Given the test results, please consider the following:
  - a. How willing would you be to move next door to Terry?
  - b. How willing would you be to spend an evening socializing with Terry?

- c. How willing would you be to make friends with Terry?
- d. How willing would you be to start working closely with Terry?
- e. How willing would you be to have Terry marry into the family?
- f. How willing would you be to have Terry as your roommate?
- g. How willing would you be to trust Terry with the care of your child?

Now, consider the condition of Jane, whose diagnosis is based on the **brain** test:

- 1. Some psychiatric disorders result in clear changes in the brain, whereas others, the effects are evident psychologically. How likely is it that Jane's condition affects her brain?
- 2. Suppose that like Terry, Jane were to give her future daughter up for adoption. Jane would have no contact with her future daughter. Given Jane's test results, how likely do you think it is that, as the daughter matures, she will exhibit the same condition as Jane?
- 3. Please now consider the following:
  - a. How treatable do you think Jane's condition will be?
  - b. How easy will it be to treat Jane's condition?
  - c. How long do you think Jane will have this condition (1 week, 1 week to 1 month, 1 month to 3 months, 3 months to 6 months, 6 months to 12 months, 1 year to 3 years, more than 3 years
- 4. Suppose you met Jane. Given the test results, please consider the following:
  - a. How willing would you be to move next door to Jane?
  - b. How willing would you be to spend an evening socializing with Jane?
  - c. How willing would you be to make friends with Jane?
  - d. How willing would you be to start working closely with Jane?
  - e. How willing would you be to have Jane marry into the family?
  - f. How willing would you be to have Jane as your roommate?
  - g. How willing would you be to trust Jane with the care of your child?

#### Schizophrenia vignette:

Nicole and Alex are two 18-year-old students who are seeking treatment because their families notice they often talk with non-existent people who they claim are invading their thoughts. Their voices have each taken on a dull, monotonous quality, and they seem to have difficulty concentrating or making sense when they speak.

A clinician suspects the two women might be suffering from schizophrenia. To better understand their condition, he administers each woman a psychological test.

The test presents the patient with various types of sounds, embedded in white noise, such that the sounds are difficult to discern. Some sounds feature human speech; other sounds feature environmental noises (i.e., the sound of a train). The test examines the patient's response to these sounds. However, the tests administered to Nicole and Alex differ on how their response is measured.

**Nicole's** response is evaluated by asking her to **press a button** to indicate whether or not the auditory stimulus contains human speech. **Alex's response** is evaluated by recording the activation of **her brain**.

Both tests are strictly normed, and the clinician knows that for people with schizophrenia, environmental sounds elicit either faster button-pressing responses or brain spikes relative to speech sounds (unlike people without schizophrenia).

The clinician has determined that Nicole and Alex's responses to the test each exhibit the abnormal pattern. But as noted for Nicole, this conclusion is based on the result of her behavioral (button-pressing) response, whereas for Alex, the result is based on her brain response (brain spike).

With this information in mind, consider these various questions about Nicole (diagnosed by the **behavioral** test).

- 1. Some psychiatric disorders result in clear changes in the brain, whereas others, the effects are evident psychologically. How likely is it that Nicole's condition affects her brain?
- 2. Suppose Nicole had an identical twin whom she never met, as her parents got divorced when they were born and the twins were raised separately. Given Nicole's test results, how likely is it that her twin sister would suffer from the same condition?
- 3. Please now consider the following:
  - a. How treatable do you think Nicole's condition will be?
  - b. How easy will it be to treat Nicole's condition?
  - c. How long do you think Nicole will have this condition (1 week, 1 week to 1 month, 1 month to 3 months, 3 months to 6 months, 6 months to 12 months, 1 year to 3 years, more than 3 years
- 4. Suppose you met Nicole. Given the test results, please consider the following:

- a. How willing would you be to move next door to Nicole?
- b. How willing would you be to spend an evening socializing with Nicole?
- c. How willing would you be to make friends with Nicole?
- d. How willing would you be to start working closely with Nicole?
- e. How willing would you be to have Nicole marry into the family?
- f. How willing would you be to have Nicole as your roommate?
- g. How willing would you be to trust Nicole with the care of your child?

Now, consider the condition of Alex, whose diagnosis is based on the **brain** test.

- 1. Some psychiatric disorders result in clear changes in the brain, whereas others, the effects are evident psychologically. How likely is it that Alex's condition affects her brain?
- 2. Suppose Alex, too, had a twin that she had never met, at the two twins were raised separately. Given Alex's test results, how likely is it that her twin sister would suffer from the same condition?
- 3. Please now consider the following:
  - a. How treatable do you think Alex's condition will be?
  - b. How easy will it be to treat Alex's condition?
  - c. How long do you think Alex will have this condition (1 week, 1 week to 1 month, 1 month to 3 months, 3 months to 6 months, 6 months to 12 months, 1 year to 3 years, more than 3 years
- 4. Suppose you met Alex. Given the test results, please consider the following:
  - a. How willing would you be to move next door to Alex?
  - b. How willing would you be to spend an evening socializing with Alex?
  - c. How willing would you be to make friends with Alex?
  - d. How willing would you be to start working closely with Alex?
  - e. How willing would you be to have Alex marry into the family?
  - f. How willing would you be to have Alex as your roommate?
  - g. How willing would you be to trust Alex with the care of your child?

#### **Bipolar Disorder vignette:**

Anna and Michelle are two 50-year-old women who are brought to the clinic by their families since they have an unrealistic belief that they are prodigious scientists and have discovered the cure for cancer. At present, they both feel omnipotent, but their families report that when this phase ends these women sink into depression, pessimism, and helplessness.

A clinician suspects the two women might be suffering from bipolar disorder. To better understand their condition, he administers each woman a psychological test.

The test presents the patient with various types of sounds, embedded in white noise, such that the sounds are difficult to discern. Some sounds feature spoken phrases – either positive ("I'm content") or omnipotent ("I'm invincible"); other sounds feature environmental noises (i.e., the sound of a train). The patient is simply asked to press a button to indicate whether or not the auditory stimulus contains human speech. However, the tests administered to Anna and Michelle differ on how their response is measured.

**Anna's** response is evaluated by asking her to **press a button** to indicate whether or not the sound corresponds to human speech. **Michelle's response** is evaluated by recording the activation of **her brain**.

Both tests are strictly normed, and the clinician knows that for people with bipolar disorder, emotionally omnipotent phrases elicit either faster button-pressing responses or brain spikes relative to emotionally positive phrases (unlike people without bipolar disorder).

The clinician has determined that Anna and Michelle's responses to the test each exhibit the abnormal pattern. But as noted for Anna, this conclusion is based on the result of her behavioral (button-pressing) response, whereas for Michelle, the result is based on her brain response (brain spike).

With this information in mind, consider various questions about Anna (diagnosed by the **behavioral** test).

- 1. Some psychiatric disorders result in clear changes in the brain, whereas others, the effects are evident psychologically. How likely is it that Anna's condition affects her brain?
- 2. Suppose Anna were to donate one of her eggs. The egg would be ultimately fertilized by a donor's sperm, and implanted in the womb of another woman who would raise the baby. This infant would never get to meet Anna. Given Anna's test results, how likely is it that, as the infant matures, she will ultimately suffer from the same condition as Anna?
- 3. Please now consider the following:
  - a. How treatable do you think Anna's condition will be?
  - b. How easy will it be to treat Anna's condition?

- c. How long do you think Anna will have this condition (1 week, 1 week to 1 month, 1 month to 3 months, 3 months to 6 months, 6 months to 12 months, 1 year to 3 years, more than 3 years
- 4. Suppose you met Anna. Given the test results, please consider the following:
  - a. How willing would you be to move next door to Anna?
  - b. How willing would you be to spend an evening socializing with Anna?
  - c. How willing would you be to make friends with Anna?
  - d. How willing would you be to start working closely with Anna?
  - e. How willing would you be to have Anna marry into the family?
  - f. How willing would you be to have Anna as your roommate?
  - g. How willing would you be to trust Anna with the care of your child?

Now, consider the condition of Michelle, whose diagnosis is based on the **brain** test.

- 1. Some psychiatric disorders result in clear changes in the brain, whereas others, the effects are evident psychologically. How likely is it that Michelle's condition affects her brain?
- 2. Suppose Michelle, too, donates an egg which ultimately is raised to be a child that she will never meet. Given Michelle's test results, how likely is it that, as the infant matures, she will ultimately suffer from the same condition as Michelle?
- 3. Please now consider the following:
  - a. How treatable do you think Michelle's condition will be?
  - b. How easy will it be to treat Michelle's condition?
  - c. How long do you think Michelle will have this condition (1 week, 1 week to 1 month, 1 month to 3 months, 3 months to 6 months, 6 months to 12 months, 1 year to 3 years, more than 3 years
- 4. Suppose you met Michelle. Given the test results, please consider the following:
  - a. How willing would you be to move next door to Michelle?
  - b. How willing would you be to spend an evening socializing with Michelle?
  - c. How willing would you be to make friends with Michelle?
  - d. How willing would you be to start working closely with Michelle?
  - e. How willing would you be to have Michelle marry into the family?
  - f. How willing would you be to have Michelle as your roommate?
  - g. How willing would you be to trust Michelle with the care of your child?