Journal of Consulting and Clinical Psychology
2017, Vol. 85, No. 11, 1052–1063

0022-006X/17/$12.00 http://dx.doi.org/10.1037/ccp0000254

Testing Positive for a Genetic Predisposition to Depression Magnifies Retrospective Memory for Depressive Symptoms

Matthew S. Lebowitz
Columbia University

Woo-kyoung Ahn
Yale University

Objective: Depression, like other mental disorders and health conditions generally, is increasingly construed as genetically based. This research sought to determine whether merely telling people that they have a genetic predisposition to depression can cause them to retroactively remember having experienced it. ***Method:*** U.S. adults (men and women) were recruited online to participate (Experiment 1: $N = 288$; Experiment 2: $N = 599$). After conducting a test disguised as genetic screening, we randomly assigned some participants to be told that they carried elevated genetic susceptibility to depression, whereas others were told that they did not carry this genetic liability or were told that they carried elevated susceptibility to a different disorder. Participants then rated their experience of depressive symptoms over the prior 2 weeks on a modified version of the Beck Depression Inventory-II. ***Results:*** Participants who were told that their genes predisposed them to depression generally reported higher levels of depressive symptomatology over the previous 2 weeks, compared to those who did not receive this feedback. ***Conclusions:*** Given the central role of self-report in psychiatric diagnosis, these findings highlight potentially harmful consequences of personalized genetic testing in mental health.

What is the public health significance of this article?
This research suggests that personalized information about genetic susceptibility to depression can distort people's recollections of having experienced depressive symptoms. This could have significant clinical implications, because currently depressive disorders (and other psychiatric disorders) are mainly diagnosed based on people's recollections of experiencing symptoms.

Keywords: depression, genetics, memory

With the rapid advance of genomic research, the public is increasingly exposed to and receptive about information concerning the genetic bases of health problems, including mental disorders (Dar-Nimrod & Heine, 2011). Within a mere 10 years (from 1996 to 2006), laypeople's genetic attributions rose from 61 to 71% for schizophrenia and from 51 to 64% for major depression (Pescosolido et al., 2010). Members of the general public have also shown a strong interest in learning about their genetic makeup. Indeed, the use of direct-to-consumer genetic testing services is so prevalent that it has generated vast databases that can now assist genomic research; using 23andMe's data from more than 300,000 customers, including more than 75,000 who reported having been diagnosed with major depression, Hyde et al. (2016) were able to identify 15 genetic loci associated with risk of major depression. It has also been projected that personalized genetic testing for susceptibility to mental disorders will likely be a widespread component of clinical practice in the near future (e.g., Couzin, 2008). As the field increasingly embraces psychiatric genetics and prepares to grapple with the use of personalized genetic information in patient care, it is imperative to understand the clinical impact that receiving such information may have.

Unfortunately, clear-cut evidence regarding the impact of personalized genetic information is sparse (Caulfield, Chandrasekharan, Joly, & Cook-Deegan, 2013), and thus it is perhaps unsurprising that direct-to-consumer genetic testing has been both celebrated as offering significant public-health benefits and also decried as potentially harmful to the public (Hogarth, Javitt, & Melzer, 2008). Personalized genetic testing may allow for personally tailored treatments, enhance diagnostic precision, and facilitate prediction of individual susceptibility to particular disorders (e.g., Thompson, Hamilton, & Hippman, 2015; Walden et al., 2015). However, genetic information could be misunderstood or misused, leading to discrimination, unnecessary treatments, and psychological distress (Drmanac, 2011).

Matthew S. Lebowitz, Center for Research on Ethical, Legal and Social Implications of Psychiatric, Neurologic and Behavioral Genetics, Department of Psychiatry, Columbia University; Woo-kyoung Ahn, Department of Psychology, Yale University.

This work was supported by grant R01-HG007653 from the National Human Genome Research Institute (National Institutes of Health). Matthew S. Lebowitz also received support from NIH Grant P50-HG007257. We thank Paul Appelbaum, Marvin Chun, John Dovidio, Avram Holmes, and Charles Sanislow for helpful suggestions, and David Guzhnay, Lily Sands, and Beth Westgate for assistance preparing study materials.

Correspondence concerning this article should be addressed to Matthew S. Lebowitz, Department of Psychiatry, Columbia University, New York State Psychiatric Institute, 1051 Riverside Drive, Unit 122, New York, NY 10032. E-mail: msl2207@cumc.columbia.edu

One well-documented public misunderstanding about genetics is the erroneous belief that genetics play a deterministic role in identity and illness (Dar-Nimrod & Heine, 2011). This misconception can cause fatalistic perceptions that disorders are immutable. For instance, the more people who attribute their own depressive symptoms to genetics and other biological factors, the more pessimistic they are about recovering in the future (Lebowitz, Ahn, & Nolen-Hoeksema, 2013).

Can people's memories of their past experience of psychiatric symptoms also be distorted due to misunderstandings about genetics? Personalized genetic information may fundamentally alter people's self-concepts (Cheung, Dar-Nimrod, & Gonsalkorale, 2014; Heine, Dar-Nimrod, Cheung, & Proulx, 2017), and such changes can bias retrospective reports of one's experiences (Robinson & Clore, 2002). Decades of cognitive research have demonstrated that "memories can be . . . distorted by current knowledge, beliefs, and expectations" (Schacter, 1999, p. 193), even leading people to recall events that never occurred (Roediger & McDermott, 1995). The process of "imaginative reconstruction or construction" (p. 213) of memories characterizes the act of remembering (Bartlett, 1932).

The present research aimed to investigate whether merely learning that one is genetically susceptible to a mental disorder could cause the person to reconstruct his or her prior experience to include more symptoms. If personalized genetic information does distort people's memories for psychiatric symptoms, the clinical consequences would likely be significant, because the current system for diagnosing most psychological disorders depends primarily on patients' retrospective reports of their experience of symptoms. For example, most symptoms used for diagnosing major depressive disorder, including depressed mood, anhedonia, and feelings of worthlessness, are assessed by subjective self-report. Thus, inflated memory for symptoms stemming from personalized genetic feedback could interfere with the accurate diagnosis of depression. Furthermore, if genetic feedback inflates people's perceptions of the extent to which they have experienced psychiatric symptoms, this could lead to self-stigma (Yen et al., 2005), resulting in various negative consequences such as low self-esteem and reduced self-efficacy (Corrigan, Larson, & Rüsch, 2009).

In studying this issue, we focused on major depression for the following reasons. Major depression is one of the most prevalent mental disorders and is a leading cause of disability worldwide (Kessler et al., 2005; Whiteford et al., 2013). In addition, despite increased attributions to genetic and other biological factors, the general public believes that major depression is also caused by socioenvironmental factors such as stress and the way a person was raised (Pescosolido et al., 2010). Thus, focusing on depression provides a particularly strong test for the effect of genetic feedback, given that genetic factors are not widely believed to be its only or most dominant cause.

To test whether receiving personalized genetic information affects recall of depressive symptoms, we conducted experimental studies using a false but credible saliva test described to participants as a proxy for gauging genetic susceptibility to depression, and manipulated the results participants received, before measuring their memory for their own depressive symptoms. In the present experiment, the deception was necessary because we sought to examine the effects of genetic feedback per se, without actual genetic differences between participants as a confound. Had we used actual genetic feedback, people with truly elevated genetic susceptibility to depression would likely score higher on self-report measures of depressive symptoms, making it impossible to distinguish the effects of receiving genetic feedback from the effects of actual genetic differences. Thus, it was necessary to randomly assign participants to the feedback they received; "effective nondeceptive alternative procedures are not feasible" (American Psychological Association, 2010). The deception lasted less than 15 min, was followed by extensive debriefing, and yielded no reports of adverse events, with study completion rates of near 100% despite an option to withdraw without losing compensation (see below for details). A similar methodology was safely used by Dar-Nimrod, Zuckerman, and Duberstein (2013) in a study that experimentally manipulated feedback participants received regarding genetic susceptibility to alcoholism to test its effect on mood and willingness to attend drinking-related workshops. All procedures detailed below were approved by the Institutional Review Board.

In two experiments, some participants were told that they were genetically susceptible to depression, while others were told that they were not. Then, participants rated their experience of depressive symptoms over the prior 2 weeks by answering items taken from the Beck Depression Inventory-II (Dozois, 2010). We hypothesized that people who were led to believe that their genes predisposed them to depression would recall having experienced more depressive symptoms than would those who were told that they did not have such a genetic predisposition.

In both experiments, we also included a condition in which participants were told that they were genetically susceptible to depression, but then viewed a video explaining that genes make only a nondeterministic contribution to the development of major depression, to counteract the genetic determinism espoused by many laypeople (Dar-Nimrod & Heine, 2011). In work by Lebowitz and Ahn (2015), this brief video successfully restored feelings of agency over depression, even among people who strongly attributed depression to biological causes (see also Lebowitz et al., 2013 for similar findings). Thus, if genetic determinism were to cause increased recall of depressive symptoms, this video intervention could have counteracted such an effect. Alternatively, if genetic information shifted people's self-concepts sufficiently to distort their recall of the past, such an effect could be more difficult to correct than fatalism about the future, hindering the effectiveness of the intervention video.

General Method for Experiments 1 and 2

Before presenting each of the experiments, this section first describes the general method used in Experiments 1 and 2. Both involved mailing "saliva testing" materials to participants in an approximate simulation of direct-to-consumer genetic testing. Participants self-administered the "saliva test" following instructions provided to them, were immediately provided with an interpretation of the "test results," and then answered questions measuring their recollection of experiencing depressive symptoms in the past.

Recruitment of Participants

Participants were recruited through Amazon.com's online Mechanical Turk (MTurk) platform—a service that allows individuals

to sign up to complete online tasks in exchange for pay (Buhrmester, Kwang, & Gosling, 2011). We first posted on MTurk an opportunity for adults in the United States to enter their mailing address in exchange for $1, noting that the actual study required materials that could not be delivered online and would thus be mailed to respondents at the address they entered. We also indicated in the same posting that participants who later completed the study itself would receive an additional $10. When participants received the study materials by mail, these were accompanied by instructions for how to complete the main portion of the study.

We collected 614 mailing addresses for Experiment 1 (which had 3 conditions), and 1,030 for Experiment 2 (which had 5 conditions). Parcels containing study materials were mailed to each address, and the response rate (i.e., the proportion of mailed parcels that resulted in an individual consenting to participate in the main study) was 75.24% in Experiment 1 and 74.56% in Experiment 2. Of those who consented to participate in the main study, the completion rate (i.e., the proportion of participants who continued through to the end of the study procedures and submitted their responses) was 95.02% in Experiment 1 and 98.18% in Experiment 2. Of these, the proportion of participants who correctly followed the instructions for the saliva test (as explained below) and could thus be included in the sample was 97.72% in Experiment 1 and 98.28% in Experiment 2. Demographics of the sample whose data were used in the results reported below are shown in Table 1.

Table 1
Demographic Characteristics of Samples in Experiments 1 and 2

Variable	Experiment 1 (N = 429)	Experiment 2 (N = 741)
Gender	52.4% female, 47.1% male, .5% no response	53.4% female, 46.3% male, .3% no response
Age	Range: 18–72 Mean: 32.52 years	Range: 18–72 Mean: 34.80 years
Education	No high school diploma: .5% High school diploma: 35.9% Associate's degree or equivalent: 16.3% Bachelor's degree or equivalent: 36.6% Postgraduate degree(s): 10.7%	No high school diploma: .2% High school diploma: 29.8% Associate's degree or equivalent: 15.2% Bachelor's degree or equivalent: 40.6% Postgraduate degree(s): 14.0%
Ethnicity	Not Hispanic or Latino: 92.1% Hispanic or Latino: 5.8% No response: 2.1%	Not Hispanic or Latino: 93.3% Hispanic or Latino: 5.4% No response: 1.3%
Race	American Indian or Alaska Native: 1.9% Asian: 7.0% Black or African American: 7.7% Native Hawaiian or Pacific Islander: .5% White: 84.1% More than one race: 2.1% Prefer not to answer: 1.4%	American Indian or Alaska Native: 1.9% Asian: 7.0% Black or African American: 6.3% Native Hawaiian or Pacific Islander: .1% White: 87.0% More than one race: 1.6% Prefer not to answer: 1.2%

Note. For race, it was possible to select more than one response.

Materials

The parcels we sent to the mailing addresses provided by MTurk users each contained a "saliva testing kit" (see Figure 1) and a letter. The saliva testing kit consisted of a hinged plastic box, inside of which was one glucose test strip (which participants were later led to believe gauged salivary levels of 5-Hydroxyindoleacetic acid as part of a genetic susceptibility test) and a small amount of mouthwash in a plastic container. Unbeknownst to participants, this mouthwash contained glucose. The letter explained how to access the web address for the main study, administered using Qualtrics.com data-collection software.

Informed Consent

When participants accessed the main portion of the study, they first viewed an onscreen informed consent document, at the bottom of which was a check-box labeled "I agree." Participants were not able to proceed unless they checked this box. We were not able to alert them in the informed consent form that there would be "genetic testing" or even "a saliva test" during the study (because an informed consent form cannot contain deception). Yet, they were told that they may experience some psychological discomfort in answering questions about their thoughts and feelings. Most importantly, they were told that they were free to skip any question or withdraw from the study at any time without losing any compensation, and thus upon first learning about the "saliva test" procedure, participants could withdraw if they did not wish to complete it.

Additional Human-Subjects Considerations

Although there was no report of adverse events in Experiment 1, we provided additional human-subjects protection in Experiment 2. After participants in Experiment 2 provided informed consent but before they proceeded to the rest of the online procedures, they were told: "If you would like to discontinue your participation in this study, it is important that you click the "Exit without completing the study" link (which can be found at the bottom of each page), instead of merely closing the browser window (exiting the study via this link will also allow you to still be compensated as a study participant)." If participants clicked this link at any point during the study, they were directed to the debriefing page. As in Experiment 1, there were no reports of adverse events in Experiment 2. In addition, this constant reminder of an option to withdraw did not increase the withdrawal rate compared to Experiment 1, suggesting that the near perfect completion rate of Experiment 1 was highly unlikely to have been due to participants' forgetting about their option to withdraw.

Saliva Test

This section describes the procedures for the "saliva test" in the three conditions that were present in both Experiments 1 and 2. After the informed consent, participants were told:

> Recent scientific research has shown that some genes can influence a person's risk of developing depression. As part of your participation in today's study, you will undergo a test to determine your genetic risk of developing "Major Depressive Disorder" (a mental disorder char-

Figure 1. Testing kit used in Experiments 1 and 2, assembled (top) and disassembled (bottom). See the online article for the color version of this figure.

> acterized by a pervasive low mood that lasts at least 2 weeks and is accompanied by low self-esteem and a loss of interest or pleasure in normally enjoyable activities). Because full-fledged DNA testing can be time-consuming, a saliva test of a chemical called 5-Hydroxyindoleacetic acid will be used. This chemical is present in many parts of the body and can be detected in saliva. Recent studies have found that if a person has genes that increase their risk of developing Major Depressive Disorder, that person will also have abnormally low levels of 5-Hydroxyindoleacetic acid in their saliva. The less 5-Hydroxyindoleacetic acid a person has in his or her saliva, the higher that person's genetic risk of Major Depressive Disorder is.

They were subsequently presented with instructions for self-administering the "saliva test." They were told to remove the mouthwash container and the test strip from the box, rinse their mouths with the mouthwash for 7 seconds (to "facilitate the 5-Hydroxyindoleacetic acid test by eliminating impurities from your saliva and increasing saliva production"), spit the mouthwash out into the empty box, insert the test strip under their tongues for 10 seconds, and then wait 30 seconds.

Participants were then asked to select the color to which the test strip had changed. In reality, the test strip was sensitive to glucose, causing it to change to brownish-green for all participants, but participants were asked to select one of three color options: "Red or Pink," "White," or "Brown or Green." This was intended to suggest that the outcome of the color change was not predetermined. Choosing a response other than Brown or Green was assumed to indicate a malfunction of the test strip or inattention to instructions; as such, participants who did not select Brown or Green were excluded from analyses (see "Recruitment of Participants" for details.). Participants who did select Brown or Green received different feedback depending on the condition to which they had been assigned.

At the end of the study, after answering optional demographic questions but before being debriefed, participants rated their agreement with the statement, "The test . . . gave accurate and reliable information about my genetic makeup" on a 5-point scale. Response options were *Strongly Disagree* (1), *Disagree* (2), *Neither Agree nor Disagree* (3), *Agree* (4), and *Strongly Agree* (5).

Debriefing

Participants read a thorough debriefing screen that informed them that no actual genetic testing was conducted in the study. For emphasis, they were told that it was very important for them to know that their genetic makeup was not tested as part of the study and that the study did not reveal any information about their genes or their risk for depression (or any other disorder). The debriefing informed them that the study's procedures were designed to lead people to believe that they were undergoing genetic testing in order to examine their psychological reactions to the test results. Participants were told that they had the right to have their data discarded if they wished after learning about the deception inherent in the study. No participant chose this option. The debriefing also informed participants that "regardless of your genetic makeup, genes alone can never make someone depressed, and depression is believed to result from a complex interaction of many different factors."

Participants were also provided with a link to an online directory for finding mental health treatment providers, the phone number to a crisis counseling hotline in case participants needed it (although they were reminded to dial 911 if they were experiencing an emergency), and contact information for the researchers and the institutional review board.

At the end of the debriefing, to verify that participants had understood its contents, all participants were required to rate three statements (e.g., "no actual genetic testing was performed on me as part of today's study") as either "true" or "false." All three statements were true in order to avoid introducing false information during the debriefing (even as part of a question). Participants were required to correctly indicate that all three statements were true in order to receive a unique "completion code" that they needed in order to be compensated for their participation (if they answered any incorrectly, they were given the opportunity to try again). Along with the completion code, additional information was displayed about the nondeterministic role of genes in depression, to help dispel any harmful misconceptions that participants might have held.

Experiment 1

Experiment 1 included three conditions: "gene-absent" (in which participants were told that they did not carry a gene predisposing them to depression), "gene-present" (in which participants were told they did carry this gene), and "gene-present/interven-

tion" (in which participants were told they carried the gene but also viewed the intervention video). We tested whether participants in the gene-present condition would report having experienced more depressive symptoms than those in the gene-absent condition, as well as whether the intervention video would effectively counteract any such difference.

Method

After the saliva test explained in General Method, participants were randomly assigned to one of the three conditions, which determined what interpretation of the test-strip's color change they would receive. Those in the gene-present and gene-present/intervention conditions were told that they carried "a gene that has been shown to significantly increase a person's risk of developing Major Depression." Those in the gene-absent condition were told that they did not carry such a gene.[1] The feedback for these conditions is presented below with the bracketed phrases showing how the two gene-present conditions (the phrase before "/") differed from the gene-absent conditions (the phrase after "/").

> The brown/green color on the sensitive region of the test strip indicates [abnormally low/normal] levels of 5-Hydroxyindoleacetic acid. This indicates that your DNA [contains/does not contain] a gene that has been shown to significantly increase a person's risk of developing Major Depression. [The following provides more information about this gene, so please read the information below carefully./The following provides more information about this gene. Although you do NOT have this gene, please read the information below carefully.] Recent scientific research has shown that a gene for something called the "serotonin transporter" is associated with a person's risk of developing Major Depression. In particular, a certain form of this gene, called the "short" form, is associated with a significantly increased risk of developing depression. This is thought to be because the "short" form of this gene can cause a chemical imbalance in the brain involving the neurotransmitter serotonin, which is important in mood. Additionally, the "short" form of the serotonin transporter gene is associated with changes in brain structure, especially in areas of the brain important for emotion, such as the amygdala. According to the National Institute of Mental Health, about 16.5% of adults in the United States will have at least one episode of Major Depression in their lifetimes. In addition, if they develop Major Depression, 32 years is the average age of onset. [People who carry the "short" form of the serotonin transporter gene, like yourself as indicated by the saliva test conducted today, are significantly more likely to have depression than people who do not carry it./People who carry the "short" form of the serotonin transporter gene are significantly more likely to have depression than people who do not carry it, like yourself as indicated by the saliva test conducted today.]

Next, all participants viewed the following background information:

> In answering the following questions, you will notice that we refer to "Major Depressive Disorder" or "an episode of Major Depression." The following explains what we mean by these terms: Everyone occasionally feels blue or sad. But these feelings are usually short-lived and pass within a couple of days. Major Depressive Disorder (also called Major Depression) is more severe than that. It refers to a low mood lasting most of the day, almost every day, for at least 2 weeks. As explained earlier, approximately 16.5% of adults in the United States appear to have an episode of Major Depression at least once in their lifetimes, and the average age of onset is 32.

Then, participants were asked two questions: "What do you think the odds are (from 0 to 100%) that you will experience an episode of Major Depression at some point in the future?" and "What do you think the odds are (from 0 to 100%) that your child or children will suffer from Major Depression at some point? (If you do not currently have children, please answer this question imagining that you have one or more children at some point in the future.)" For each item, participants selected a point on a 0–100 scale to indicate their response. These two questions were not the main focus of the current study, which was designed to examine the effect of genetic feedback on retrospective memory for depressive symptoms. However, these items were included because past research has demonstrated that people who believe that they experience depression because of a biological predisposition are more likely to expect that depression will be unavoidable in the future (Lebowitz et al., 2013; Kemp, Lickel, & Deacon, 2014). Thus, we reasoned that if our manipulation of genetic feedback was effective, it would likely have a significant effect on these ratings. Indeed, the ratings were highly correlated with one another, $r = .75$, $p < .001$, and the average of the two was significantly higher in the gene-present conditions[2] ($M = 51.80\%$) than in the gene-absent condition ($M = 30.47\%$), $t(427) = 8.05$, $p < .001$, $d = .77$.

At this point, participants in the gene-present/intervention condition were taken to the page containing the intervention video explaining that genes do not deterministically cause major depression. It discussed the notion that genes alone are not sufficient to cause depression, provided information about gene-by-environment interactions and epigenetics, and discussed actions that people can take to combat or decrease their risk of depression even if they are genetically predisposed (Lebowitz & Ahn, 2015). The video is available at https://www.youtube-nocookie.com/embed/hupQ_kkJXrg. After watching the video, participants were asked to "write a few sentences about the information you learned from the video you just watched . . . and give at least a few examples of how you (or somebody else) might use the information you learned to prevent or overcome depression." Instructing participants to compose these reflections took advantage of the so-called "saying is believing" effect, in which people come to internalize a viewpoint more strongly after they have advocated for it themselves (Higgins, 1999; Lebowitz & Ahn, 2015; Lebowitz et al., 2013; Walton & Cohen, 2011). As detailed in the Appendix, analyses of the written reflections showed that 93.6% of participants demonstrated full comprehension of the video.

Next, to test the effect of genetic feedback on retrospective recall of depression, which was our primary outcome of interest, two measures were taken. First, all participants were asked, "Have you had at least one episode of Major Depression (a low mood lasting most of the day, almost every day for at least 2 weeks) in your life?" To answer this question, they moved a slider along a 0–4 scale, labeled *Absolutely No* (0), *Probably No* (1), *Uncertain*

[1] Although no single gene causes major depression, the "saliva test results" referred to a single gene for methodological simplicity.

[2] We collapsed ratings from the two gene-present conditions for this analysis because their procedures had been identical up to this point.

(2), *Probably Yes* (3), and *Absolutely Yes* (4). Three participants did not provide a response.

Second, all participants completed the Beck Depression Inventory-II (BDI-II), a well-validated and widely used measure of depression (Dozois, 2010), on which participants rated the extent to which they had experienced various depressive symptoms over the past 2 weeks (higher scores indicate more symptomatology). For instance, under "sadness," they chose among 0 (*I do not feel sad*), 1 (*I feel sad much of the time*), 2 (*I am sad all the time*), and 3 (*I am so sad or unhappy that I can't stand it*). As another example, in the "concentration difficulty" item, they were to choose among 0 (*I can concentrate as well as ever*), 1 (*I can't concentrate as well as usual*), 2 (*It's hard to keep my mind on anything for very long*), and 3 (*I find I can't concentrate on anything*). We omitted one item, "Suicidal Thoughts or Wishes," because our online procedures precluded appropriate responses to reports of suicidality; thus, participants rated 20 items, with possible scores ranging from 0 to 60.

Results

We first examined perceived credibility of the saliva test. A one-way ANOVA revealed no significant effect of condition on ratings of the saliva test's credibility, $F(2, 426) = 1.08$, $p = .342$, indicating that perceived credibility of the test was not a confound. On average, the test was seen as credible: credibility ratings were significantly higher than the scale midpoint of 3 in the gene-absent condition ($M = 3.59$, $t(142) = 48.16$, $p = .001$, $d = .66$), the gene-present condition ($M = 3.62$, $t(144) = 45.17$, $p < .001$, $d = .64$), and the gene-present/intervention condition ($M = 3.74$, $t(140) = 44.66$, $p < .001$, $d = .74$).

We then examined one of the main dependent measures of the study, participants' ratings of whether they had experienced at least one episode of major depression. A one-way ANOVA revealed a significant main effect of condition on these ratings, $F(2, 423) = 7.50$, $p = .001$. Pairwise comparisons revealed that compared to participants in the gene-absent condition ($n = 143$, $M = 2.00$, 95% C.I. [1.73, 2.26]), those in the gene-present condition ($n = 143$, $M = 2.68$, 95% C.I. [2.44, 2.93]) gave significantly higher ratings, $p = .001$ using Tukey's "honestly significant difference" (HSD) test and $p < .001$ using one-sided Dunnet t tests treating the gene-absent condition as a control condition ($d = .44$). Furthermore, even participants in the gene-present/intervention condition ($n = 140$, $M = 2.50$, 95% C.I. [2.24, 2.54]) gave significantly higher ratings than those in the gene-absent condition, $p = .017$ using Tukey's HSD test and $p = .006$ using one-sided Dunnet t tests treating the gene-absent condition as a control condition ($d = .32$).

Next, we examined BDI-II Scores. A one-way ANOVA revealed a significant main effect of condition, $F(2, 426) = 6.71$, $p = .001$ (see Figure 2). Compared to participants in the gene-absent condition ($n = 143$ $M = 11.09$, 95% C.I. [9.36, 12.82]), significantly higher BDI-II scores were observed among those in the gene-present condition ($n = 145$, $M = 16.04$, 95% C.I. [14.03, 18.06]), $p = .001$ using Tukey's HSD test and $p < .001$ using one-sided Dunnet t tests treating the gene-absent condition as a control condition ($d = .44$). BDI-II scores in the gene-absent condition were also higher than those in the gene-present/intervention condition ($n = 141$, $M = 13.91$, 95% C.I. [11.97, 15.86]), $p = .098$ using Tukey's HSD test and $p = .036$ using one-sided Dunnet t tests treating the gene-absent condition as a control condition ($d = .26$). See the Appendix for further analyses of all dependent measures excluding participants who did not demonstrate full understanding of the intervention video.

In order to provide a rough framework for interpreting the levels of depression recalled by participants in each condition, we conducted one-sample t tests comparing BDI-II scores in each condi-

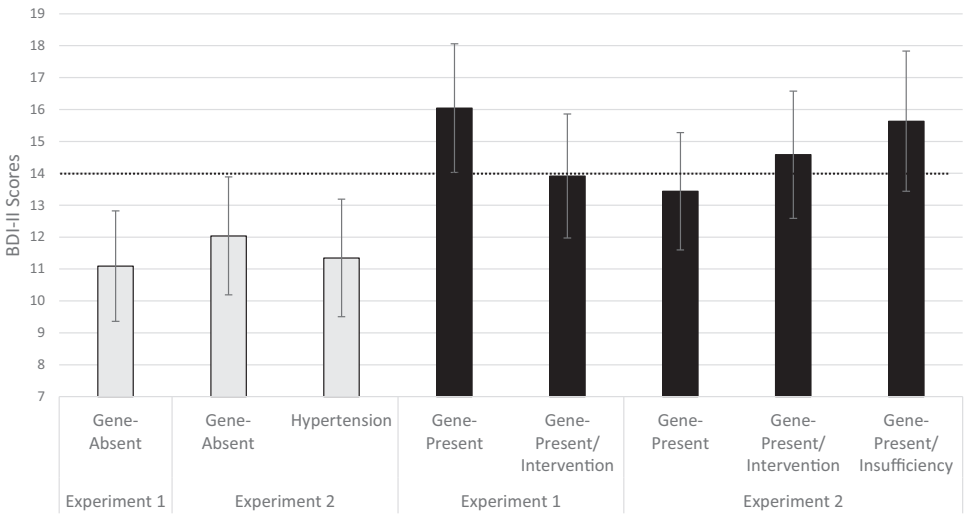

Figure 2. BDI-II scores by condition in Experiments 1 and 2. Possible BDI-II scores range from 0–60, with higher scores indicating more self-reported depression over the prior 2 weeks. Gray bars indicate conditions in which participants were not told that they were genetically predisposed to depression; black bars indicate conditions in which participants were told they were genetically predisposed to depression. Error bars represent 95% confidence intervals. Dotted line represents a score of 14, the cutoff for "mild" depression (Dozois, 2010).

tion against 14, which is conventionally recognized as a cutoff score for at least "mild" depression (Dozois, 2010). Participants in the gene-absent condition scored significantly lower than 14, $t(142) = 3.33$, $p = .001$, $d = .28$, while those in the gene-present condition scored significantly higher than 14, $t(144) = 2.00$, $p = .047$, $d = .17$. Those in the gene-present/intervention condition did not differ significantly from 14, $t(140) = .09$, $p = .931$, $d = .01$.

Discussion

Although randomly assigned to conditions, participants who were told that they were genetically predisposed to depression reported having experienced more severe depressive symptoms than those told they did not have such elevated genetic susceptibility. The BDI-II scores of those in the gene-absent condition suggested the absence of depression, whereas those of participants in the gene-present condition suggested the presence of depression.

An intervention video explaining that genes do not deterministically cause depression failed to eliminate this effect, even though nearly all participants demonstrated understanding of the main message of the video. As reported in the Appendix, even when analyses were limited to participants who evinced full comprehension of the video, the intervention did not have significant benefits. Experiment 2 further attempted to replicate these effects.

Experiment 2

Experiment 2 addressed several limitations of Experiment 1. First, it is unclear whether the higher BDI scores in the gene-present conditions were due simply to negative emotion induced by "bad news," or a reaction to feedback about depression susceptibility specifically. Thus, Experiment 2 included a hypertension condition, wherein the saliva test purportedly revealed elevated genetic susceptibility to hypertension rather than depression. Because participants in the hypertension condition also received "bad news," if that were the source of effect, their BDI-II scores would be higher—as in the gene-present conditions—than those in the gene-absent condition.

Second, it is unclear whether the difference found in Experiment 1 stemmed from the gene-present feedback inflating participants' recall of experiencing depression, the gene-absent feedback suppressing participants' recall of experiencing depression, or both. Experiment 2's hypertension condition, wherein participants received no feedback about their susceptibility to depression, served as a baseline condition to distinguish these possibilities. For instance, if the hypertension condition yielded BDI-II scores lower than the gene-present condition, but not different from the gene-absent condition, this would suggest that the difference in BDI-II scores between the gene-absent and gene-present condition were attributable to the gene-present feedback increasing participants' BDI-II scores rather than the gene-absent feedback decreasing participants' BDI-II scores.

Third, in Experiment 1 the intervention did not provide a significant benefit in counteracting the negative effects of the gene-present feedback; this may have occurred because it was too future-oriented, emphasizing strategies to counteract one's genetic susceptibility to depression. Because one cannot change the past, this intervention might not have been highly effective in preventing changes in one's memory for the past. Thus, we included a "gene-present/insufficiency" condition, in which we simply stated that the gene targeted by the saliva test was not sufficient to guarantee that one would experience depression.

Method

Experiment 2 used the same saliva test as Experiment 1. There were five conditions: hypertension, gene-absent, gene-present, gene-present/insufficiency, and gene-present/intervention. The gene-absent, gene-present, and gene-present/intervention conditions were identical to those in Experiment 1.

Participants in the hypertension condition received the same instructions about the saliva test as described in General Method, except that "Major Depressive Disorder" and "depression" were replaced with "hypertension (high blood pressure)" in the first instance, and "hypertension" thereafter. They were also told that hypertension is "a condition in which the long-term force of the blood against artery walls is high enough that it may cause health problems." After the saliva test, participants in the hypertension condition received feedback similar to the gene-present condition except that "Major Depression" was replaced with "hypertension (high blood pressure)" and that the explanation for the gene was:

> Recent scientific research has shown that a gene for something called the "angiotensinogen receptor" is associated with a person's risk of developing hypertension. In particular, a certain form of this gene, called the "C" form, is associated with a significantly increased risk of developing hypertension. This is thought to be because the "C" form of this gene can make the body more sensitive to proteins that trigger the narrowing of blood vessels. This narrowing of blood vessels, in turn, can cause increases in blood pressure. According to the Centers for Disease Control, about 16.5% of adults in the United States have hypertension. In addition, if they develop hypertension, 32 years is the average age of onset. People who carry the "C" form of the angiotensinogen receptor gene, like yourself, as indicated by the saliva test conducted today, are significantly more likely to have hypertension than people who do not carry it.

The prevalence and age-of-onset information in the feedback used in the hypertension condition was written to mirror the information provided about Major Depressive Disorder in the other conditions, to eliminate the possibility that this information would be a confound.

The gene-present/insufficiency condition was the same as the gene-present condition, except that participants were additionally told, "Having the 'short' form of the serotonin transporter gene does not mean that a person definitely develops Major Depression. At this point, however, scientists do not know the exact likelihood."

Experiment 2 omitted the ratings of participants' expectations about the likelihood that they or their children would experience depression in the future, as well as the item about past experiences of depression used in Experiment 1. This change was intended to ensure that any observed effects on BDI-II scores were not attributable merely to first having completed such ratings, whose validity has not been well established like that of the BDI-II.

All participants then completed the BDI-II and ended the experiment as explained earlier. During the debriefing phase, we also sought to explicitly address the possibility that the study's impact on self-perception might persist despite the presence of a debriefing—a possibility known as a *perseverance effect* (Ross, Lepper, & Hubbard, 1975)—by using so-called "process debriefing" in which participants were explicitly warned about the perseverance effect and advised to be vigilant about it (McFarland, Cheam, & Buehler, 2007).

Results

We first examined the perceived credibility of the saliva test. A one-way ANOVA revealed that there was no significant difference across conditions in ratings of the saliva test's credibility, $F(4, 736) = .75$, $p = .555$. The saliva test was again largely perceived as credible. The mean credibility rating was significantly higher than the scale midpoint of three in each of the five conditions (i.e., the gene-absent condition, $M = 3.69$, $t(159) = 51.51$, $p < .001$, $d = .76$; the gene-present condition, $M = 3.62$, $t(141) = 44.67$, $p < .001$, $d = .64$; the gene-present/intervention condition, $M = 3.61$, $t(151) = 46.42$, $d = .63$; the gene-present/insufficiency condition, $M = 3.52$, $t(138) = 44.35$, $p < .001$, $d = .56$; and the hypertension condition, $M = 3.54$, $t(146) = 45.46$, $p < .001$, $d = .57$).

We initially compared the BDI-II scores of the three gene-present conditions to examine whether the intervention video or the genetic insufficiency information moderated the impact of learning that one is genetically predisposed to depression (see Figure 2). A one-way ANOVA comparing the BDI-II scores of the gene-present condition ($n = 142$, $M = 13.44$, 95% C.I. [11.60, 15.28]), the gene-present/intervention condition ($n = 153$, $M = 14.58$, 95% C.I. [12.59, 16.57]), and the gene-present/insufficiency condition ($n = 139$, $M = 15.63$, 95% C.I. [13.44, 17.83]) revealed no significant effect of condition, $F(2, 431) = 1.13$, $p = .324$. This indicated that the intervention video and genetic insufficiency information had no effect, so we collapsed data across the three gene-present conditions for the next analyses. (As reported in the Appendix, excluding those who did not demonstrate full comprehension of the intervention video did not change the pattern of the results.)

We conducted a one-way ANOVA comparing BDI-II scores in the gene-present conditions, the gene-absent condition, and the hypertension condition, which revealed a significant main effect of condition (see Figure 2), $F(2, 738) = 5.17$, $p = .006$. Follow-up comparisons revealed that the gene-present conditions ($n = 434$, $M = 14.54$, 95% C.I. [13.39, 15.70]) yielded significantly higher BDI-II scores than the gene-absent condition ($n = 160$, $M = 12.04$, 95% C.I. [10.19, 13.89]), $t(592) = 2.23$, $p = .026$, $d = .18$, as well as significantly higher BDI-II scores than the hypertension condition ($n = 147$, $M = 11.35$, 95% C.I. [9.50, 13.19]), $t(579) = 2.79$, $p = .005$, $d = .23$. BDI-II scores in the hypertension condition were not significantly different from those in the gene-absent condition, $t(305) = .52$, $p = .602$, $d = .06$.

Even though the three gene-present conditions were not significantly different from each other, we also conducted Tukey's HSD tests and one-sided Dunnet t tests to provide more detailed analyses. Each gene-present condition was compared to each of the two control conditions (i.e., the gene-absent condition and the hypertension condition). The results of these analyses are reported in Table 2. In general, while these analyses continued to suggest that BDI-II scores were higher in the gene-present/insufficiency condition than in the gene-absent or hypertension conditions, this effect was not consistently significant with regard to the other two gene-present conditions.

This suggested a possible discrepancy between Experiments 1 and 2, in which pairwise analyses comparing BDI-II scores in the gene-absent condition against those in the gene-present conditions more consistently remained significant when adjusting for multiple comparisons in the former than in the latter. One possible explanation we considered for such a discrepancy was that the failure of some such comparisons to reach in Study 2 could have stemmed from Type II error. Specifically, in the present research we included in analyses all participants regardless of their ratings of the saliva test's credibility (i.e., even those who rated the saliva test as not credible).[3] As a result, our analyses were quite conservative in that they included even participants who explicitly indicated that they did not believe the results of the saliva test, potentially diluting the effects of our manipulation and causing our analyses to underestimate the magnitude of the effects of genetic feedback. Considering this, we felt that the risk of Type II error may have been substantial.

In response to these considerations, we conducted a mega-analysis in which we included all participants from both experiments who were in one of the three conditions that followed identical procedures in the two experiments: gene-absent, gene-present, and gene-present/intervention. By combining participants across the two experiments, this analysis benefitted from increased sample size, which we hoped would help to decrease the risk of Type II error. A one-way ANOVA comparing the gene-absent, gene-present, and gene-present/intervention conditions (combining participants from the two experiments) revealed a significant main effect of condition, $F(2, 881) = 6.166$, $p = .002$. A pairwise comparison revealed that across the two experiments, participants in the gene-absent condition ($n = 303$, $M = 11.69$, 95% C.I. [10.40, 12.98]) had significantly lower BDI-II scores than those in the gene-present condition ($n = 287$, $M = 14.90$, 95% C.I. [13.51, 16.29]), $p = .003$ from Tukey's HSD test, $p = .001$ from one-sided Dunnet t test ($d = .28$). Participants in the gene-absent condition also had significantly lower BDI-II scores than those in the gene-present/intervention condition ($n = 294$, $M = 14.26$, 95% C.I. [12.88, 15.65]), $p = .022$ from Tukey's HSD test, $p = .008$ from one-sided Dunnet t test ($d = .22$). The mega-analysis revealed no significant difference between the gene-present and gene-present/intervention conditions ($p = .792$ from Tukey's HSD test).

Finally, as in Experiment 1, we conducted one-sample t tests comparing BDI-II scores in each condition of Experiment 2 against 14 to provide a rough guide for interpreting the levels of depression recalled by participants in each condition. The results showed that mean BDI-II scores were significantly lower than 14 in the gene-absent condition, $t(159) = 2.10$, $p = .038$, $d = .17$, and in the hypertension condition $t(146) = 2.85$, $p = .005$, $d = .23$. However, this was not the case in the gene-present condition, $t(141) = .61$, $p = .546$, $d = .05$, the gene-present/insufficiency condition, $t(138) = 1.47$, $p = .144$, $d = .12$, or the gene-present/intervention condition, $t(152) = .58$, $p = .565$, $d = .05$.

[3] We chose not to limit our analyses to only participants who found the saliva test credible, even though doing so would have yielded substantially more robust effects of the gene-present feedback than what is reported here. This is because such analyses would have introduced a potential confound for the following reasons. Participants who were more depressed at baseline might have been more likely to rate the saliva test as credible in the gene-present conditions, and participants who were less depressed at baseline might have been more likely to rate the test as credible in the gene-absent conditions. Thus, if we had retained only participants who rated the saliva test as credible, there would have been no way of knowing whether any observed difference in recall of depressive symptoms stemmed merely from baseline differences in depressive symptoms, or from a combination of baseline differences in depression and some effect of our experimental manipulations. By including all participants in our analyses, we sought to remove any baseline differences in preexisting depression levels across the experimental conditions while retaining the effect of our experimental manipulations.

Table 2
P-Values of Pairwise Comparisons Using Tukey's HSD Tests (P-Value Before Slash) and One-Sided Dunnet T-Tests Treating the Gene-Absent or the Hypertension Condition as a "Control" Condition and Comparing Each Gene-Present Condition to Them (P-Value After Slash)

Comparator	Gene-present	Gene-present/intervention	Gene-present/insufficiency
Gene-absent	.849/.374 (d = .12)	.330/.094 (d = .21)	.073/.017 (d = .29)
Hypertension	.574/.190 (d = .19)	.134/.033 (d = .27)	.022/.005 (d = .35)

Note. Each cell shows the p-values of comparisons between the condition shown in the first column of that row and the condition shown in the first row of that column. Values in parentheses are the effect sizes for each comparison.

Discussion

In Experiment 2, an analysis that combined all participants who received the gene-present feedback suggested that they reported having experienced more depressive symptoms than those who received the gene-absent or hypertension feedback. This was confirmed by a mega-analysis combining data from both experiments. Overall, these results appear to support the main conclusion of Experiment 1: although the information participants received about their genetic predispositions was determined at random, BDI-II scores were significantly higher among those told that they were genetically predisposed to depression than those not given such feedback.

Experiment 2 also eliminated several alternative explanations for Experiment 1's findings. BDI-II scores in Experiment 2's hypertension condition were not higher than those in the gene-absent condition, suggesting that merely receiving any bad news cannot account for elevated BDI-II scores in the gene-present conditions, as participants in the hypertension condition also received bad news. Moreover, the gene-present/insufficiency condition was not significantly different from the other two gene-present conditions, demonstrating that even explicitly denying genetic determinism without orienting participants toward the future failed to counteract the effect of the gene-present feedback. In fact, the pairwise comparisons reported in Table 2 showed that it was actually the gene-present/insufficiency condition that showed the most reliable effect of genetic feedback on retrospective symptom recall.

Additionally, in Experiment 1, it was unclear whether the difference between the gene-absent and the gene-present conditions stemmed from the gene-present feedback inflating participants' recall of experiencing depression, the gene-absent feedback suppressing participants' recall of experiencing depression, or both. Participants in the hypertension condition of Experiment 2 did not receive any gene-absent feedback regarding depression, so the pattern of results suggests that telling people they were genetically predisposed to depression inflated their BDI-II scores.

General Discussion

Previously, information about biological susceptibility to mental disorders was shown to affect future-oriented psychological variables (e.g., confidence in controlling one's own symptoms; Dar-Nimrod et al., 2013; Kemp et al., 2014). The current studies present novel results suggesting that feedback indicating elevated genetic susceptibility to depression can lead people to retrospectively report having experienced more depressive symptoms.

This is particularly alarming given the current psychiatric diagnostic system's heavy reliance on retrospective self-report of symptoms, not only for depression but also for most mental disorders. Thus, the current results present potential challenges for the field to grapple with as psychiatric genetic testing becomes more prevalent in the future, by suggesting that delivering personalized genetic information to patients may complicate the process of making psychiatric diagnoses.

Moreover, in the present work, an approach that has been shown to mitigate genetic attributions' detrimental effects on people's feelings of agency in confronting their symptoms (Lebowitz & Ahn, 2015; Lebowitz et al., 2013) failed to counteract the effects of genetic feedback on symptom recall. This may reflect the fact that overall, the widespread bias toward assuming that genes affect health deterministically is difficult to counteract; indeed, much of the existing work that has effectively weakened such assumptions has done so by disproportionately emphasizing the role of nongenetic factors, which may be viewed as more controllable (Heine et al., 2017). Such approaches, like the intervention tested in the present research, may be effective at increasing people's feelings of control over their health in the future. However, this does not guarantee that they would be particularly effective at mitigating the effects of genetic feedback on symptom recall, especially if the latter reflect a deep-seated change in self-concept.

Taken together, these findings have important implications for practicing mental health clinicians, including psychologists. Although genetic tests are not yet widely used among mental health practitioners, increasing usage of diagnostic and predictive genetic testing in psychiatry is considered likely, and there is value in understanding and preparing for the complications they may entail before they become more widespread (Appelbaum & Benston, 2017). Growing numbers of clinicians may wish to incorporate genetic test results to inform the care they provide, and even those who do not may increasingly encounter patients who have already obtained results through direct-to-consumer genetic testing services. For instance, the Food and Drug Administration recently granted permission to one company to provide consumers with personalized information about their genetic susceptibility to disorders such as Parkinson's disease, Alzheimer's disease, and Celiac disease (Food & Drug Administration, 2017); direct-to-consumer genetic test purportedly revealing psychiatric vulnerabilities are not unrealistic in the near future. At the very least, given our results, clinicians may wish to elicit a patient's self-reported symptom history before the patient receives any genetic information, although it is an empirical question whether this would lessen subsequent memory distortions caused by genetic

feedback. In cases wherein patients already believe themselves to be genetically predisposed to major depression, the current results suggest that clinicians may wish to seek out information beyond self-reported symptoms (e.g., corroboration from a patient's friends and family members who are not aware of the patient's genetic predisposition) to confirm their diagnoses.

One limitation of the present research is that it only examined short-term effects of personalized genetic feedback, as the deception about the authenticity of the saliva test could not ethically be extended to examine long-term effects. While some research suggests that short-term negative effects of medical test results can dissipate over time (Shaw, Abrams, & Marteau, 1999), we know of no study to date that has examined how personalized genetic information affects recall of the past in the long run. Even if effects like increases in health anxiety after the return of test results do tend to decrease with time, altered perceptions of the past may be more robust, as suggested by their seeming imperviousness to our video intervention.

Another potential limitation of the present studies is that to minimize the duration of the deception, we used the BDI-II, rather than structured diagnostic interviews, which are currently the "gold standard" for formal psychiatric diagnosis. Nonetheless, given the well-established validity of the BDI-II (e.g., Dozois, 2010), it likely serves as a suitable proxy.

Our findings also raise a number of significant questions for future research. For instance, it is not clear whether personalized genetic susceptibility information would affect symptom recall for other conditions besides depression. We speculate that the answer may depend on how genetically determined the disorder in question is perceived to be in the first place. Indeed, we found such an effect with depression in spite of the public's general tendency to attribute depressive symptoms more strongly to psychosocial and environmental factors than to biological ones (Angermeyer & Dietrich, 2006). The effect could be even more pronounced in other disorders that are seen as more genetically based, and potentially less pronounced in cases wherein genes are seen as less determinative.

Additionally, it remains to be determined whether the effects we observed in the present research are unique to personalized genetic feedback. For example, the present findings cannot speak to the question of whether individuals who were told that their past experiences predisposed them to depression would react with elevated recall of depressive symptoms, just as people told they are genetically predisposed to depression appear to do. The present research focused specifically on the effects of genetic feedback, as patients' individual genetic information is expected to be increasingly incorporated to guide clinical care (Guttmacher, McGuire, Ponder, & Stefánsson, 2010). Given numerous demonstrations of top-down effects on memory (e.g., Schacter, 1999), it is quite possible that any individualized information about risk factors may lead to distortions like those observed in the present research.

Another important question for future research concerns whether the effect of genetic information found in the current research would depend on individual differences in levels of depression. Because the present work used BDI-II scores as a dependent variable, nothing is known about participants' baseline levels of depressive symptoms. On one hand, those who are currently depressed may be more likely to be affected by the genetic feedback because these people may be more fatalistic than those who are less depressed. Indeed, depressed individuals are vulnerable to becoming preoccupied by negative information (Gotlib & Joormann, 2010), especially if the information is self-relevant (Beck & Clark, 1988). Thus, they may be at particular risk of psychological harm from learning negative information about their genes. As a result, the genetic feedback could lead to even more inflated symptom recall among those who are already suffering from depression. On the other hand, indications that one is genetically susceptible to depression may be more novel and surprising to those who are not suffering from depression, resulting in greater impact on such people. In general, while there is no clear reason to assume that the present findings would not generalize to a more depressed clinical sample—and indeed, there is evidence suggesting that some people with symptoms of depression do believe that biological factors play a deterministic role in causing their symptoms (Lebowitz et al., 2013)—future research could specifically examine this issue.

In conclusion, the current results indicate that caution is warranted as personalized health-related genetic test results become increasingly available in the health care system. Overall, our findings add to the growing evidence that genetic attributions can create a powerful filter through which people view their own mental health (Dar-Nimrod et al., 2013; Haslam, 2011; Kemp et al., 2014; Lebowitz, 2014), with clinically important consequences.

References

American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author. Retrieved from http://www.apa.org/ethics/code/principles.pdf

Angermeyer, M. C., & Dietrich, S. (2006). Public beliefs about and attitudes towards people with mental illness: A review of population studies. *Acta Psychiatrica Scandinavica, 113,* 163–179. http://dx.doi.org/10.1111/j.1600-0447.2005.00699.x

Appelbaum, P. S., & Benston, S. (2017). Anticipating the ethical challenges of psychiatric genetic testing. *Current Psychiatry Reports, 19,* 39. http://dx.doi.org/10.1007/s11920-017-0790-x

Bartlett, F. C. (1932). *Remembering*. New York, NY: Cambridge University Press.

Beck, A. T., & Clark, D. A. (1988). Anxiety and depression: An information processing perspective. *Anxiety Research, 1,* 23–36. http://dx.doi.org/10.1080/10615808808248218

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3–5. http://dx.doi.org/10.1177/1745691610393980

Caulfield, T., Chandrasekharan, S., Joly, Y., & Cook-Deegan, R. (2013). Harm, hype and evidence: ELSI research and policy guidance. *Genome Medicine, 5,* 21. http://dx.doi.org/10.1186/gm425

Cheung, B. Y., Dar-Nimrod, I., & Gonsalkorale, K. (2014). Am I my genes? Perceived genetic etiology, intrapersonal processes, and health. *Social and Personality Psychology Compass, 8,* 626–637. http://dx.doi.org/10.1111/spc3.12138

Corrigan, P. W., Larson, J. E., & Rüsch, N. (2009). Self-stigma and the "why try" effect: Impact on life goals and evidence-based practices. *World Psychiatry, 8,* 75–81. http://dx.doi.org/10.1002/j.2051-5545.2009.tb00218.x

Couzin, J. (2008). Science and commerce. Gene tests for psychiatric risk polarize researchers. *Science, 319,* 274–277. http://dx.doi.org/10.1126/science.319.5861.274

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin, 137,* 800–818. http://dx.doi.org/10.1037/a0021860

Dar-Nimrod, I., Zuckerman, M., & Duberstein, P. R. (2013). The effects of learning about one's own genetic susceptibility to alcoholism: A randomized experiment. *Genetics in Medicine, 15,* 132–138. http://dx.doi.org/10.1038/gim.2012.111

Dozois, D. J. A. (2010). Beck Depression Inventory-II. In I. B. Weiner & W. E. Craighead (Eds.), *The Corsini encyclopedia of psychology* (4th ed., pp. 210–211). New York, NY: Wiley. http://dx.doi.org/10.1002/9780470479216.corpsy0113

Drmanac, R. (2011). The advent of personal genome sequencing. *Genetics in Medicine, 13,* 188–190.

Food & Drug Administration. (2017, April 6). *FDA allows marketing of first direct-to-consumer tests that provide genetic risk information for certain conditions* (Press release). Retrieved from https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm551185.htm

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: Current status and future directions. *Annual Review of Clinical Psychology, 6,* 285–312. http://dx.doi.org/10.1146/annurev.clinpsy.121208.131305

Guttmacher, A. E., McGuire, A. L., Ponder, B., & Stefánsson, K. (2010). Personalized genomic information: Preparing for the future of genetic medicine. *Nature Reviews Genetics, 11,* 161–165. http://dx.doi.org/10.1038/nrg2735

Haslam, N. (2011). Genetic essentialism, neuroessentialism, and stigma: Commentary on Dar-Nimrod and Heine (2011). *Psychological Bulletin, 137,* 819–824. http://dx.doi.org/10.1037/a0022386

Heine, S. J., Dar-Nimrod, I., Cheung, B. Y., & Proulx, T. (2017). Essentially biased: Why people are fatalistic about genes. *Advances in Experimental Social Psychology, 55,* 137–192. http://dx.doi.org/10.1016/bs.aesp.2016.10.003

Higgins, E. T. (1999). "Saying is believing" effects: When sharing reality about something biases knowledge and evaluations. In J. M. Levine, D. M. Messick, & L. L. Thompson (Eds.), *Shared cognition in organizations: The management of knowledge* (1st ed., pp. 33–49). Mahwah, NJ: Erlbaum.

Hogarth, S., Javitt, G., & Melzer, D. (2008). The current landscape for direct-to-consumer genetic testing: Legal, ethical, and policy issues. *Annual Review of Genomics and Human Genetics, 9,* 161–182. http://dx.doi.org/10.1146/annurev.genom.9.081307.164319

Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., . . . Winslow, A. R. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics, 48,* 1031–1036. http://dx.doi.org/10.1038/ng.3623

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. http://dx.doi.org/10.1016/j.brat.2014.02.009

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62,* 593–602. http://dx.doi.org/10.1001/archpsyc.62.6.593

Lebowitz, M. S. (2014). Biological conceptualizations of mental disorders among affected individuals: A review of correlates and consequences. *Clinical Psychology: Science and Practice, 21,* 67–83. http://dx.doi.org/10.1111/cpsp.12056

Lebowitz, M. S., & Ahn, W. K. (2015). Emphasizing malleability in the biology of depression: Durable effects on perceived agency and prognostic pessimism. *Behaviour Research and Therapy, 71,* 125–130. http://dx.doi.org/10.1016/j.brat.2015.06.005

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,* 518–527. http://dx.doi.org/10.1037/a0031730

McFarland, C., Cheam, A., & Buehler, R. (2007). The perseverance effect in the debriefing paradigm: Replication and extension. *Journal of Experimental Social Psychology, 43,* 233–240. http://dx.doi.org/10.1016/j.jesp.2006.01.010

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. *American Journal of Psychiatry, 167,* 1321–1330. http://dx.doi.org/10.1176/appi.ajp.2010.09121743

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128,* 934–960. http://dx.doi.org/10.1037/0033-2909.128.6.934

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814. http://dx.doi.org/10.1037/0278-7393.21.4.803

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32,* 880–892. http://dx.doi.org/10.1037/0022-3514.32.5.880

Schacter, D. L. (1999). The seven sins of memory. Insights from psychology and cognitive neuroscience. *American Psychologist, 54,* 182–203. http://dx.doi.org/10.1037/0003-066X.54.3.182

Shaw, C., Abrams, K., & Marteau, T. M. (1999). Psychological impact of predicting individuals' risks of illness: A systematic review. *Social Science & Medicine, 49,* 1571–1598. http://dx.doi.org/10.1016/S0277-9536(99)00244-0

Thompson, C., Hamilton, S. P., & Hippman, C. (2015). Psychiatrist attitudes towards pharmacogenetic testing, direct-to-consumer genetic testing, and integrating genetic counseling into psychiatric patient care. *Psychiatry Research, 226,* 68–72. http://dx.doi.org/10.1016/j.psychres.2014.11.044

Walden, L. M., Brandl, E. J., Changasi, A., Sturgess, J. E., Soibel, A., Notario, J. F. D., . . . Müller, D. J. (2015). Physicians' opinions following pharmacogenetic testing for psychotropic medication. *Psychiatry Research, 229,* 913–918. http://dx.doi.org/10.1016/j.psychres.2015.07.032

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science, 331,* 1447–1451. http://dx.doi.org/10.1126/science.1198364

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., . . . Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: Findings from the Global Burden of Disease Study 2010. *The Lancet, 382,* 1575–1586. http://dx.doi.org/10.1016/s0140-6736(13)61611-6

Yen, C. F., Chen, C. C., Lee, Y., Tang, T. C., Yen, J. Y., & Ko, C. H. (2005). Self-stigma and its correlates among outpatients with depressive disorders. *Psychiatric Services, 56,* 599–601. http://dx.doi.org/10.1176/appi.ps.56.5.599

Appendix

Analyses of Reflections Written by Participants in the Gene-Present/Intervention Condition

The analyses reported in the main text suggested that the intervention video was not effective. To examine the possibility that this may have occurred because some participants did not fully understand the video, we examined the text written by participants in the gene-present/intervention conditions of both experiments. For this analysis, we coded each written reflection using a 3-point scale. We assigned a rating of 3 to reflections that demonstrated full comprehension of the main message of the intervention (i.e., that genes do not deterministically cause depression). We assigned a rating of 2 to reflections that demonstrated partial comprehension (e.g., those that mentioned that environmental or controllable factors could also play a role in depression but did not explicitly mention the nondeterministic nature of genes' involvement). Finally, we assigned a rating of 1 to reflections that demonstrated minimal comprehension or did not demonstrate comprehension, as well as to a small number of reflections that demonstrated comprehension of the intervention's message but then disavowed its persuasiveness. Both authors initially read and rated all reflections independently. The two authors' initial ratings were in agreement for 95.04% of participants in Experiment 1 and 93.46% of participants in Experiment 2. For cases wherein the two initial ratings did not agree, the authors discussed their reasoning and arrived at a final rating by consensus.

Experiment 1

In the gene-present/intervention condition of Experiment 1, 1.4% of reflections received a rating of 1, 5.0% received a rating of 2, and 93.6% received a rating of 3. To assess whether the intervention might have been more effective among participants whose written reflection suggested comprehension of its message, we repeated the primary analyses reported in the main text including only participants whose reflections were assigned a rating of 3. These participants had a mean score of 2.50 (n = 131, 95% C.I. [2.23, 2.77]) on the measure of whether they had experienced at least one episode of major depression, and had a mean BDI-II score of 13.53 (n = 132, 95% C.I. [11.56, 15.50]). One-way ANOVAs revealed a significant effect of condition on participants' ratings of whether they had experienced at least one episode of major depression, $F(2, 414) = 7.47$, $p = .001$, and on BDI-II scores, $F(2, 417) = 6.76$, $p = .001$.

A pairwise comparison of the gene-present/intervention condition to the gene-absent condition revealed that participants in the gene-present/intervention condition whose reflections were assigned a score of 3 were still significantly more likely to recall having experienced at least one episode of major depression in the past, $t(272) = 2.63$, $p = .009$, $d = .32$. However, unlike in the analysis reported in the main text, their BDI-II scores were only marginally, rather than significantly, higher, $t(273) = 1.85$, $p = .066$, $d = .22$.

Next, we compared the gene-present and gene-present/intervention conditions. When all gene-present/intervention participants were included, the two conditions did not differ significantly in BDI-II scores, $t(284) = 1.50$, $p = .135$, $d = .18$, or participants' ratings of whether they had experienced at least one episode of major depression, $t(281) = .99$, $p = .323$, $d = .12$. However, when gene-present/intervention participants were limited to those whose reflections received a rating of 3, a marginal difference in BDI-II scores emerged, $t(275) = 1.75$, $p = .081$, $d = .21$. The difference in past-depressive-episode ratings remained nonsignificant, $t(272) = .99$, $p = .323$, $d = .12$. While the marginal difference in BDI-II scores may suggest that participants who fully understood and internalized the message of the intervention video were buffered from some of the negative effects of the gene-present feedback to some small extent, there was no evidence for such a protective effect in Experiment 2 (see below).

Experiment 2

In the gene-present/intervention condition of Experiment 2, 2.6% of reflections received a rating of 1, 9.2% received a rating of 2, and 88.2% received a rating of 3. To assess whether the intervention might have been more effective among participants whose written reflection suggested comprehension of its message, we repeated the primary analyses reported in the main text including only participants in the gene-present/intervention condition whose reflections were assigned a rating of 3. These participants had a mean BDI-II score of 14.82 (n = 135, 95% C.I. [12.71, 16.93]), compared to a mean of 14.58 (as reported in the main text) when no participants were excluded. This indicates that the intervention was not more beneficial among participants whose reflections received a rating of 3. Indeed, as in the analyses reported in the main text, a one-way ANOVA comparing the BDI-II scores of the three gene-present conditions revealed no significant effect of condition, $F(2, 413) = 1.16$, $p = .313$. We also conducted a one-way ANOVA comparing BDI-II scores in the gene-present conditions, the gene-absent condition, and the hypertension condition, which revealed a significant main effect of condition, $F(2, 720) = 5.38$, $p = .005$. Follow-up comparisons revealed that BDI-II scores in the gene-present conditions ($n = 416$, $M = 14.62$, 95% C.I. [13.44, 15.80]) were significantly higher compared to those in the gene-absent condition, $t(574) = 2.29$, $p = .022$, $d = .19$, as well as compared to those in the hypertension condition, $t(561) = 2.86$, $p = .005$, $d = .24$. These findings mirror those reported in the main text.

Received March 9, 2017
Revision received August 9, 2017
Accepted August 11, 2017 ■