Are Symptom Clusters Explanatory?
A Study in Mental Disorders and Non-Causal Explanation

Daniel A. Wilkenfeld (daniel.wilkenfeld@berkeley.edu)*
Jennifer Asselin (asselin.7@osu.edu)**
Tania Lombrozo (lombrozo@berkeley.edu)*
*Department of Psychology, UC Berkeley
**Department of Philosophy, The Ohio State University

Abstract

Three experiments investigate whether and why people accept explanations for symptoms that appeal to mental disorders, such as: "She experiences delusions because she has schizophrenia." Such explanations are potentially puzzling, as mental disorder diagnoses are made on the basis of symptoms, and the DSM implicitly rejects a commitment to some common, underlying cause. Do laypeople nonetheless conceptualize mental disorder classifications in causal terms? Or is this an instance of non-causal explanation? Experiment 1 shows that such explanations are indeed found explanatory. Experiment 2 presents participants with novel disorders that are stipulated to involve or not involve an underlying cause across symptoms and people. Disorder classifications are found more explanatory when a causal basis is stipulated, or when participants infer that one is present (even after it's denied in the text). Finally, Experiment 3 finds that merely having a principled, but non-causal, basis for defining symptom clusters is insufficient to reach the explanatory potential of categories with a stipulated common cause. We discuss the implications for accounts of explanation and for psychiatry.

Keywords: explanation, understanding, mental disorders

Anecdotal evidence suggests that people consider diagnostic categories to be explanatory. For instance, one might explain a patient's high blood sugar levels by appeal to diabetes. This is also the case for mental disorders. A blurb about the film *A Beautiful Mind* claims that "the film displays the idea that Nash is a genius *because he has schizophrenia*" (Covell, 2013, emphasis added). In the wake of a mass shooting, it is common for people to cite the shooter's mental illness in explaining the atrocity (Craghill & Clement, 2015).

Consider the most basic form of such explanatory claims: those that appeal to a diagnostic category to explain the presence of a symptom. "He suffers from hallucinations because he has schizophrenia." "She is persistently sad because she has depression." Such claims may be common, but they're also potentially puzzling. The *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, known as the DSM-5, catalogues recognized mental disorders along with criteria for diagnosis. Importantly, these criteria involve the presence, duration, and severity of various *symptoms*, and are explicit that it is these symptoms which define the disorder: "A mental disorder is a syndrome *characterized* by clinically significant disturbance of an individual's cognition, emotion regulation, or behavior…" (American Psychiatric Association, 2013, p. 20, emphasis added). Explaining a symptom by appeal to a diagnostic category thus borders on a tautology – it's (almost) like explaining that someone has little money *because he is poor* (that's just what it *means* to be poor).

What, then, is the value of these explanations? One possibility is that laypeople's beliefs about mental disorders depart from the stipulations of the DSM – as even experts' do when it comes to causal relationships between symptoms (Kim & Ahn, 2002a, 2002b). Rather than thinking of mental disorders in terms of symptom clusters, laypeople may treat diagnostic labels more like medical disease labels, which often pick out some underlying condition that's *causally responsible* for observable symptoms. Past work finds that while clinicians generally reject the idea that mental disorders share causal "essences" (Ahn, Flanagan, Marsh, & Sanislow, 2006), laypeople do not (Cooper & Marsh, 2015). On this view, then, mental disorder classifications could be explanatory because they (perhaps indirectly) offer *causal* explanations. But another possibility is that laypeople accept at least some *non-causal* explanations (e.g., Prasada & Dillingham, 2006), and that DSM diagnostic categories support them.

Despite these connections to prior work, no research – to our knowledge – has investigated whether and why mental disorder classifications are found explanatory. Given their explicitly non-causal basis, DSM categories provide fruitful terrain in which to explore questions about the relationship between explanation and causation. Is a disease classification only explanatory when it picks out some common, underlying causal structure? Or can mere symptom clusters – clusters that are not tied by causal etiology – explain the presence of the symptoms associated with the corresponding diagnosis? We explore these questions in three experiments.

Experiment 1

In Experiment 1 we verify a presupposition of our project: that laypeople do, in fact, find mental disorder categories explanatory. To do so, we assess people's willingness to accept claims of the form: "Alex experiences hallucinations because she has schizophrenia." In addition to looking at absolute levels of agreement with such statements, we include medical diseases and non-explanations for comparison.

Methods

Participants Fifty-three adults (23 male, 30 female, mean age = 34) were recruited through the Amazon Mechanical Turk marketplace (MTurk) and participated in exchange for monetary compensation. In all experiments, participation was restricted to users with an IP address within the United States and an approval rating of at least 95% based on at least 50 previous tasks. An additional seven participants were excluded prior to analysis for failing to consent, failing to complete the experiment, or giving an incorrect response to one of the reading comprehension or attention check questions (detailed below). For all experiments, those who had completed another experiment in this line of research were ineligible.

Materials and Procedures Participants were presented with 12 vignettes in random order. In each vignette, participants were told that a character displayed a symptom, went to a doctor, and was correctly diagnosed with a disease that had five listed symptoms. They were also told the character was suffering from a second, unrelated problem. For example, participants read:

> Rachel…has been experiencing a decreased range of motion (in her fingers)....She was correctly diagnosed with arthritis. Symptoms of arthritis include decreased range of motion, pain, swelling, stiffness, and inflammation.
>
> As it happens, Rachel also has headaches, which are not a symptom of arthritis.

After each vignette, participants were asked two reading comprehension questions, which asked both whether the main symptom (e.g., "decreased range of motion") and the unrelated symptom (e.g., "headaches") were symptoms of the disorder mentioned—anyone who answered either question incorrectly for any vignette was excluded from further analysis. (Participants were also excluded if they failed an attention check based on Oppenheimer, Meyvis, and Davidenko, (2009), at the end of the task.)

Participants were then asked to evaluate an explanation claim on a 7-point Likert scale from "Strongly Disagree" (1) to "Strongly Agree" (7):

> [Name] has [symptom] *because* s/he has [disease].
> (e.g., Rachel has a decreased range of motion *because* she has arthritis.)

In eight of the cases – four based on medical conditions ("Medical") and four based on mental disorders ("Mental") – the symptom asked about was in fact the first one listed for the disorder (e.g., whether someone had a skin rash because of measles or hallucinations because of schizophrenia). In the remaining four cases ("Control") – two medical and two mental – the symptom was from the unrelated condition (e.g., "Rachel has headaches because she has arthritis"). The control items were included to ensure that participants did not provide indiscriminately high ratings.

Results & Discussion

Responses to the "because" statements were averaged for each participant into three sets: those for the medical, mental, and control vignettes. These average ratings were then analyzed with a repeated-measures ANOVA with vignette type as a within-subjects factor, revealing a significant effect, $F(2, 104) = 571.644$, $p < .001$, $\eta_p^2 = .917$. Bonferroni pairwise comparisons revealed no significant difference ($p = .071$) between the Medical ($M = 6.35$, $SD = .611$) and Mental ($M = 6.14$, $SD = .779$) disorders, but that both differed significantly (ps $< .001$) from control cases ($M = 1.87$, $SD = .759$). Moreover, the ratings for Mental disorders were significantly above the scale midpoint ($p < .001$), suggesting that participants indeed found references to mental disorder categories explanatory.

Experiment 2

Having established that mental disorder classifications are found explanatory, we consider (in Experiment 2) whether this depends upon the causal structure of the category. Specifically, we varied whether diagnosis was based on the presence of a common cause or on a cluster of symptoms. For comparison, we also included a condition in which the disease was diagnosed in an arbitrary fashion. To limit effects of prior knowledge about mental disorders, we introduced fictional disorders on an alien planet.

Methods

Participants Participants were 141 adults (63 male, 77 female, 1 other, mean age 32) who were recruited through Amazon Mechanical Turk and participated in exchange for monetary compensation. Sixty-seven participants were excluded prior to analysis based on the same criteria as Experiment 1.

Materials Participants were presented with a vignette in which an alien was diagnosed with a mental disorder. Participants were divided into three groups ("Condition"). One group was told the diagnosis was made based on the presence of a particular cause ("*stipulated cause,"* N = 34), a second group was presented with diagnoses based on symptom clusters ("*symptom,"* N = 53), and a third group was presented with diagnoses based on which disorder name was pulled out of a hat ("*random draw*," N = 54).[1] Participants received symptoms corresponding to one of three disorders, which were modeled on depression, schizophrenia, or borderline personality disorder. For example, participants in the symptom/depression group read:

> John is an alien on the planet Zorg. Lately he has been experiencing a number of troubling symptoms, including persistent sadness.

[1] An equal number of participants was included in each condition prior to exclusions; however, participants in the *cause* condition did significantly worse ($p < .001$) on a true/false question about whether the disorder was diagnosed on the basis of common symptoms (this was supposed to be false for them). There were no other significant differences on reading comprehension or attention checks (ps > .240).

> Recently, John went to the doctor to find out what was wrong. The doctor consulted her medical textbook, which is used by all doctors on Zorg as the standard for defining and diagnosing illness. ***It says that a doctor should diagnose a given mental illness when and only when the patient exhibits some number (but not necessarily all) of the symptoms on a list of symptoms corresponding to that illness.***
>
> For example, it says that Gordon’s Disease should be diagnosed when a patient has some number (but not necessarily all) of the symptoms of persistent sadness, trouble falling asleep, difficulty maintaining a stable weight, light-headedness, and difficulty concentrating. The book is very clear that ***the disease always has some of these symptoms.*** However, it doesn’t always have the same cause in different people, or even the same cause for all symptoms within a given person. For example, it could be caused by a virus in some people, but by a genetic predisposition in others. Or even for the same person, some symptoms could be caused by a virus, and others by a genetic predisposition***. All that matters for having the disease is having the right set of symptoms***.
>
> John does in fact have a number of these symptoms, so the doctor diagnoses him as having Gordon’s Disease.[2]

All participants then rated their agreement with three statements about the disorder’s explanatory status (in random order, not labeled for participants) from 1 (“Strongly Disagree”) to 7 (“Strongly Agree”):

> *Because*: [Name] is [symptom] *because* he has [disorder].
>
> *Understand*: I *understand* why [name] is [symptom].
>
> *Token Cause*: [Disorder] is the *cause* of [name’s] [symptom].

Participants were also asked to generalize properties across individuals with the same diagnosis, but in the interest of space, we do not report these results here.

After these agreement and generalization ratings, participants were asked to answer the following question about the basis for the disorder, rated on a 7-point scale:

> *Inferred Common Cause*: How likely do you think it is that there is a common cause behind all cases of [disorder]?

Participants were excluded from analyses if they could not correctly answer any of three true/false questions regarding the basis on which the disorder was diagnosed. For example, they had to answer whether “Gordon’s Disease is diagnosed on the basis of symptoms sharing a particular cause.” (The correct answers varied by condition.) At the end of the task, participants were asked what real disorder they thought was closest to the one in the vignette. Only 38 of 141 participants correctly identified the disorder with which they were presented, and whether people identified the disorder had no impact on other responses. Finally, participants supplied standard demographic information, reported any problems with the survey, and had the same attention check used in Experiment 1.

Results & Discussion

Responses were analyzed in a mixed ANOVA with condition (3: *random draw*, *symptom*, *stipulated cause*) and disorder (3: *depression*, *schizophrenia*, *borderline personality*) as between-subjects factors, and statement (3: *because*, *understand*, *token cause*) as a within-subjects factor. This analysis revealed a significant main effect of condition, $F(2, 132) = 25.80$, $p < .001$, $\eta_p^2 = .281$. Tukey post-hoc tests indicated significantly higher ratings for *stipulated cause* than *symptom* ($p < .001$), which were in turn significantly higher than *random draw* ($p < .005$) (See Figure 1). There were no other significant effects. Because statement did not interact with condition, we averaged the three responses to create a single “explanation score”.

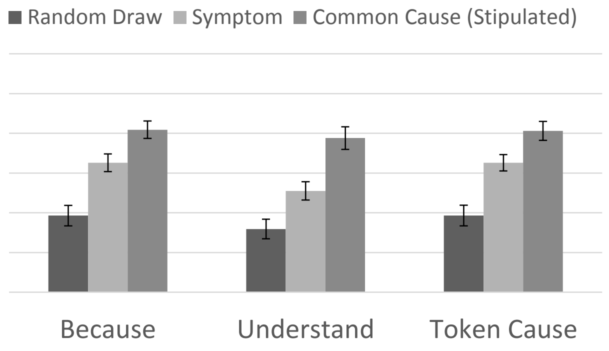

Figure 1: Experiment 2 Mean DVs by Condition. Error Bars ± 1 SEM

These results suggest that whether a diagnostic category is considered explanatory depends, at least in part, on the basis for category membership. The results also suggest that sharing a common cause is not necessary for conferring some explanatory potential: participants in the *symptom* condition gave significantly higher ratings than those in the *random draw* condition, although their ratings did not differ from the scale mid-point ($p = .890$.) However, it could be that these participants assumed the presence of a common cause, even though the vignette stipulated in those cases that none was present.

[2] We used the locution that a person “has” a disorder. Reynaert and Gelman (2007) found systematic differences in beliefs about disorder permanence depending on whether a noun-phrase, adjective-phrase, or possessive-phrase (“has”) was used, which could also affect explanation judgments. This could not, however, account for differences we find *across* our conditions. Reynaert and Gelman (2007) cite the *Publication Manual* of the American Psychological Association as suggesting it’s best to frame disorders in terms of possessive phrases (e.g., “has Gordon’s Disease”).

To investigate whether this occurred, we first examined responses to *inferred common cause* as a function of condition using a one-way ANOVA with *condition* as a between-subjects factor and *inferred common cause* as a dependent variable (see Figure 2). This analysis revealed a significant main effect, $F(2, 138) = 34.26$, $p < .001$, $\eta_p^2 = .332$; Tukey post-hoc tests revealed that all differences were highly significant (*p*s < .001).

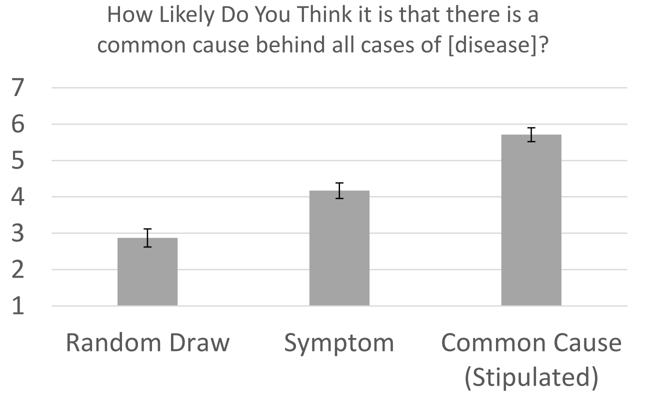

Figure 2: Experiment 2 Mean Inferred Common Cause ratings by Condition, from low (1) to high (7) likelihood. Error Bars ± 1 SEM

Although inferences to a common cause varied significantly across conditions, mean ratings reveal that many participants in the *symptom* condition (and even some in the *random draw* condition) believed there was a reasonable probability that one existed. Could it be, then, that diagnostic categories were only found explanatory to the extent that people believed the category picked out a common cause, regardless of the condition to which they were experimentally assigned?

To test this possibility, we ran a hierarchical regression. An initial model used the variable *inferred common cause* to predict explanation scores (see Figure 3). This model was highly significant ($p < .001$), with $R^2 = .35$. A second model that also included condition, coded as two dichotomous variables, accounted for more variance, with $R^2 = .40$ (this increase in R^2 was significant as assessed by a significant change in F-scores, $p < .001$). In this second model, *inferred common cause* had an unstandardized coefficient of .38 ($p < .001$), being in the stipulated cause condition (yes=1, no=0) had an unstandardized coefficient of .41 ($p = .184$), and being in the random condition (yes=1, no=0) had an unstandardized coefficient of -.72 ($p = .008$).

In sum, people's assessments of explanations were affected by *both* their inferences regarding the existence of a common cause and by the experimental manipulation of diagnostic procedure. This suggests that some factor (or factors) other than the presence of a common cause – and that varied across conditions – played a role in modulating judgments. In Experiment 3, we examined whether a relevant difference between the *common cause (stipulated)* condition and the *symptom* condition was the fact that in the former case, the symptom cluster had a non-arbitrary basis (the common cause) and presumably practical implications for treatment.

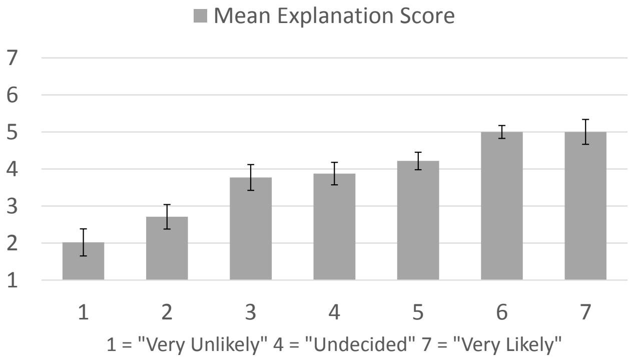

Figure 3: Experiment 2 Mean Explanation Score as a Function of Response to *Inferred Common Cause.*

Experiment 3

In this experiment, participants considered mental disorder classifications based on symptom clusters (without an underlying common cause), but in one case, the symptom cluster was motivated by non-arbitrary practical considerations, and in the second case, it was described as the result of historical accident. If having some principled basis for defining a symptom cluster – even when it's not causal – is sufficient to support good explanations, we would expect higher explanation ratings in the former case than in the latter.

Method

Participants Ninety-two adults (43 male, 49 female, 1 other, mean age 38) were recruited through MTurk and participated in exchange for monetary compensation. Fifty-nine participants were excluded prior to analysis based on the same criteria used in Experiments 1 and 2.

Materials Participants were presented with a single vignette in which an alien was diagnosed with a mental disorder, where the disorder was characterized by always sharing some cluster of the same symptoms. Participants were divided into two experimental groups ("Condition"). In the *reason* condition (N = 47), participants were told that the symptoms were grouped together into one disorder on the bases of treatment and prognosis. In the *no reason* condition (N = 45), participants were told that the symptoms were grouped together by historical accident. Both conditions stressed that there was no shared causal mechanism behind all instances of the disorder. As in Experiment 2, the vignettes were based on one of three mental disorders from the DSM: depression, schizophrenia, or borderline personality disorder.

In the *reason* condition involving depression, for example, participants read:

> Even though the symptoms of Gordon's Disease do not share a common cause, they were grouped together for principled reasons. When these symptoms occur in conjunction, they interfere with multiple facets of life in a way that interferes with close personal relationships and

> with daily routines, making it difficult for people to obtain the social support and physical well being that could facilitate treatment. As a result, people who exhibit clusters of these symptoms are at heightened risk for suicide, and also need to pursue alternative treatment plans. It's because patients with these symptoms experience similar risks and are best suited to particular treatments that they're grouped under a common disorder.

For participants in the *no reason* condition, the relevant paragraph was replaced by the following:

> The symptoms of Gordon's Disease were grouped together under one disorder by historical accident, rather than for any principled reason. Had 18th century doctors documented the associated symptoms in a different order, the disease might have been defined by a very different cluster of symptoms.

All participants were then asked to rate their agreement with the measures used in Experiment 2, and also completed comprehension and attention checks. As in Experiment 2, only a minority of participants (33 out of 92) were able to identify the real disorder, and whether or not the disorder was correctly identified had no impact on responses.

Results & Discussion

Responses were analyzed in a mixed ANOVA with condition (2: *reason*, *no reason*) and disorder (3: *depression*, *schizophrenia*, *borderline personality*) as between-subjects factors, and statement (3: *because*, *understand*, *token cause*) as a within-subjects factor (See Figure 4). This analysis did not find a significant effect of condition (p = .830): overall, mean responses to the three statements were the same across the reason (M = 3.809, SD = 1.391) and no reason (M = 3.882, SD = 1.409) conditions. However, there was a significant interaction between condition and statement, (p = .004): only "understand" ratings decreased numerically from the *reason* to the *no reason* condition, though this decrease was not itself significant (p = .195), nor were the numerical increases found for "because" (p = .530) and "token cause" (p = .170). So while the significant interaction is suggestive and merits further scrutiny, we collapsed the three ratings into a single explanation score (as in Experiment 2) for subsequent analyses.

As in Experiment 2, there was a reliable association between explanation scores and the probability assigned to *inferred common cause*, r = .281, p = .007. Interestingly, responses to *inferred common cause* did not themselves vary across conditions (p = .888).

In sum, when evaluating whether a diagnostic category offers an explanation for its symptoms, participants were insensitive to the question of whether the symptoms were grouped on the basis of a (non-causal) principled reason or a historical accident. However, once again, there was a reliable association between explanation ratings and inferences concerning the existence of a common cause.

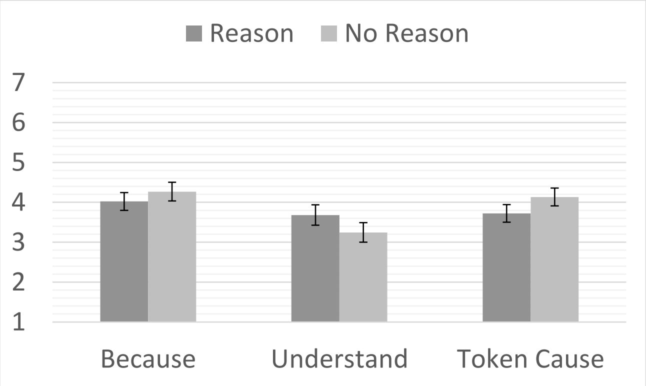

Figure 4: Experiment 3 Mean DVs by Condition. Error Bars ± 1 SEM

General Discussion

Across three experiments, we find that people are willing to explain symptoms by appeal to mental disorder classifications, but that their willingness to do so depends in large part on causal assumptions that are not endorsed by the DSM, and indeed denied in some of our vignettes. Specifically, Experiment 1 used real mental disorder classifications, and found that these classifications were not only found explanatory, but as explanatory as medical diagnoses. Experiment 2 used fictional disorder names in an alien context, which allowed us to stipulate the causal basis for each disorder classification. We found that classifications were most explanatory when they corresponded to a common cause. However, participants also found classifications based on symptom clusters significantly more explanatory than those based on a random draw. This pattern of results was largely, but not exclusively, driven by participants' assumptions about the presence of a common cause: even when the characterization of the disorder denied a common cause, participants in the symptom condition often inferred that one existed. Finally, Experiment 3 found that a non-arbitrary but non-causal basis for grouping symptoms was insufficient to improve the explanatory potential of a symptom cluster.

One robust finding from Experiments 2 and 3 is that people consider diagnostic categories more explanatory when they correspond to a common cause across cases. This is broadly consistent with causal accounts of explanation in philosophy (e.g., Woodward, 2003), according to which explanations identify one or more causes of what's being explained. However, causal accounts of explanation are both more and less demanding. On the one hand, a causal explanation need not identify a *single* common cause. It's typically sufficient to pick out a cause (or causes) that operated in the case being explained. It's not entirely clear, though, what this means when the explanation invokes a *diagnostic category* with some causal basis rather than the causes themselves. It could be that people take the diagnostic category to pick out a circumscribed *set* of causes, and thus find it explanatory by virtue of its implicit causal content, even when this content falls short of identifying a single common cause.

In another sense, causal accounts of explanation may be insufficiently demanding: inferring a common cause in our *symptom* conditions was typically enough to boost explanation ratings just above the scale mid-point, but it did not subsume the effects of experimental condition. It could be that a diagnostic category is more explanatory when it is itself *characterized* in causal terms, which is a condition that goes beyond the mere existence of a common cause.

Another possibility is that (some of) our participants were engaged in a genuinely non-causal form of explanation. Indeed, while the dominant approach to explanation within philosophy is causal, there are a variety of alternatives (Woodward, 2014), including those that focus on unification (e.g., Friedman, 1974), pragmatic import (e.g., Wilkenfeld, 2014), or argumentative structure (e.g., Hempel, 1965). Similarly, some psychological approaches to explanation suggest more formal accounts (e.g., Prasada & Dillingham, 2006).

While it's certainly possible that participants endorsed mental disorder classifications as explanatory for some non-causal reason, the findings from Experiment 3 speak against the most obvious possibilities. In particular, introducing a non-arbitrary basis for the symptom cluster, as we did in the *reason* condition, should have made the diagnostic categories better candidates for explanation on most accounts: the manipulation made the symptoms more inferentially useful, introduced pragmatic import, and arguably suggested some basis for unification. There's clearly more work to be done, but our initial findings present a puzzle for non-causal approaches.

How would clinicians respond in our task? Given prior work suggesting that expertise is associated with weaker beliefs in causal essences underlying mental disorders (Ahn et al., 2006; Cooper & Marsh, 2015), we speculate that clinicians would be less likely to endorse the explanations offered here. Similarly, we speculate that if laypeople knew what mental disorders were really thought to be by clinicians (Cooper & Marsh, 2015), they would not consider them (as) explanatory. However, our findings also suggest that dislodging laypeople's causal-essentialist beliefs about mental disorder classifications is no easy task. In the *symptom* condition of Experiment 2, and in both conditions of Experiment 3, many participants believed a common cause was likely, even though it had been explicitly denied.

While psychiatrists are largely interested in a classification scheme that best serves diagnosis and treatment, rather than one that reflects lay intuitions, the present results might have important implications. There's evidence that treatments are less effective if the patient does not believe the treatment will work (e.g., Seligman 1991), and people's intuitive beliefs about mental disorders have implications for their views about the efficacy of different kinds of treatment. It follows that lay beliefs could inform psychiatric practice.

The impact of our results on philosophy is more pronounced. People's persistence in seeing causes where there are (by stipulation) none to be had—and the fact that that tendency seems to account for much of their explanatory judgments—underscores the central role of causation in explanation.

References

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

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders: DSM-5. Washington, D.C: American Psychiatric Association.

Cooper, J. A., & Marsh, J. K. (2015). The influence of expertise on essence beliefs for mental and medical disorder categories. *Cognition, 144*, 67-75.

Covell, K. (2013, May 23). A beautiful mind [Presentation]. Retrived from https://prezi.com/vfuqvvu3pzxq/a-beautiful-mind/

Craighill, P., & Clement, S. (2015, October 26). What Americans blame most for mass shootings. *Washington Post*. Retrieved from http://tinyurl.com/oj79jrw

Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy, 71*(1), 5-19.

Harvey, A.G., Soehner, A., Lombrozo, T., Bélanger, L., Rifkin, J. & Morin, C.M. (2013). 'Folk theories' about the causes and treatment of insomnia. Cognitive Research and Therapy, 37, 1048-1057. doi:10.1007/s10608-013-9543-2

Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science by Carl G. Hempel*. New York: Free Press.

Kim, N. S., & Ahn, W. (2002a). The influence of naive causal theories on lay concepts of mental illness. American Journal of Psychology, 115(1), 33-66.

Kim, N. S., & Ahn, W. (2002b). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. Journal of Experimental Psychology: General, 131(4), 451.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867-872.

Prasada, S., & Dillingham, E. M. (2006). Principled and statistical connections in common sense conception. *Cognition, 99*(1), 73-112.

Reynaert, C. C., & Gelman, S. A. (2007). The influence of language form and conventional wording on judgments of illness. Journal of Psycholinguistic Research, 36(4), 273-295.

Seligman, M. E. (1991). *Learned optimism*. New York: Knopf.

Strevens, M. (2004). The causal and unification approaches to explanation unified—causally. Noûs, 38(1), 154-176.

Wilkenfeld, D. A. (2014) Functional explaining: A new approach to the philosophy of explanation. *Synthese, 191*(14), 3367-3391.

Woodward, J. (2003). *Making things happen: A theory of causal explanation* Oxford University Press.

Woodward, J. (2014). Scientific explanation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*.