

2016

Illusion of Understanding in a Misunderstood Field

Andrew Stephen Zeveney
Lehigh University

Follow this and additional works at: <http://preserve.lehigh.edu/etd>



Part of the [Psychology Commons](#)

Recommended Citation

Zeveney, Andrew Stephen, "Illusion of Understanding in a Misunderstood Field" (2016). *Theses and Dissertations*. 2902.
<http://preserve.lehigh.edu/etd/2902>

This Thesis is brought to you for free and open access by Lehigh Preserve. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Lehigh Preserve. For more information, please contact preserve@lehigh.edu.

Illusion of Understanding in a Misunderstood Field

by

Andrew S. Zeveney

A Thesis

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

Psychology

Lehigh University

May 23rd, 2016

Copyright 2016
Andrew Zeveney

Master's Thesis Signature Sheet

Thesis is accepted and approved in partial fulfillment of the requirements for the Master of Science in Psychology.

Illusion of Understanding in a Misunderstood Field
Andrew Zeveney

Date Approved

Jessecae Marsh
Advisor and Committee Chair

Almut Hupbach
Committee Member

Barbara Malt
Committee Member

Gordan Moskowitz
Department Chair

TABLE OF CONTENTS

List of Figures	v
List of Tables	vi
Abstract	1
Introduction	2
Overview of Experiments	11
Experiment 1	11
Experiment 2	15
Experiment 3	25
Experiment 4	33
General Discussion	38
References	44
Appendix A	49
Vita	51

LIST OF FIGURES

Figure 1: Personal understanding by time – Experiment 2.	19
Figure 2: Personal understanding across condition – Experiment 3.	28
Figure 3: Personal understanding by disorder cause – Experiment 3.	30
Figure 4: Personal understanding across condition – Experiment 4.	36

LIST OF TABLES

Table 1: Example stimuli in each domain – Experiment 1.	13
Table 2: Mean society/gap ratings by condition – Experiment 1.	14
Table 3: Mean society/gap ratings by condition – Experiment 2.	21
Table 4: Gap rating correlations by domain – Experiment 2.	22
Table 5: Society rating correlations by domain – Experiment 2.	22
Table 6: Mean manipulation ratings by condition – Experiment 4.	35

Abstract

Humans fail to fully understand the world around them and to recognize their limited understanding. The illusion of explanatory depth (IOED) exemplifies these failures: people believe they understand the world more than they actually do and only realize the illusory nature of this belief when they attempt to explain phenomena. An unexplored factor of the IOED is how people may become overconfident by confusing their own understanding with others' understanding. In four experiments, I compared the IOED in devices, where it is typically examined, with mental health, a domain where society has a more limited understanding. In Experiments 1 and 2, I demonstrate that laypeople believe society understands mental health less than devices and that people demonstrate a smaller IOED in mental health than in devices. Experiment 3 shows that explanation is necessary for the illusion to be revealed in mental health. Finally, Experiment 4 suggests that explicitly describing others' understanding as limited eliminates the illusion. Implications for meta-cognition and for mental health are discussed.

Introduction

Humans encounter a vast number of phenomena on a daily basis but only possess a shallow level of understanding of most of these phenomena (Wilson & Keil, 1998). In addition to a limited understanding of many everyday domains, people lack an understanding of their own understanding and tend to believe that they are much more skilled in a variety of domains than they actually are (Dunning, Johnson, Ehrlinger, & Kruger, 2003). In this thesis, I specifically explore the way that people fail to understand their own understanding by examining the illusion of explanatory depth (Rozenblit & Keil, 2002), people's belief that they can explain a phenomenon better than they are actually able to. I first discuss explanations in general and then findings related to people's illusory beliefs that they understand more than they do. I examine how these illusory beliefs may be supported by people's propensity to confuse what they understand about a phenomenon with what other people understand about that phenomenon. I focus specifically on the mental health domain, a field where understanding is currently very limited.

Explanations and the illusion of explanatory depth

Explanations are attempts to increase understanding of a phenomenon by illuminating several aspects of that phenomenon (Kitcher, 1981; Wilson & Keil, 1998). An explanation for a phenomenon almost always involves descriptions of the causes (e.g., Einhorn & Hogarth, 1986) and in particular, reveals the complex system of mechanisms that contributes to the phenomenon (Bechtel & Abrahamsen, 2005; Glennan, 2002). People have an innate drive to explain phenomena in the world (Gopnik, 2000; Kosslyn, 1995), with explanation being described as central to human understanding and

communication (Lombrozo, 2006). Explanation serves as a particularly important driving force in the field of science, where researchers seek to generate an explanation for why or how a phenomenon happens that accurately reflects the reality of that phenomenon (e.g., Salmon, 1989; Strevens, 2008). Overall, explanations serve an important role in human reasoning.

Explanations can be useful in helping people to evaluate their own comprehension. When people attempt to generate an explanation for a phenomenon, they not only learn what they know but also become aware of “gaps” in their understanding: those parts of the explanation that are difficult or impossible to generate (Keil, Rozenblit, & Mills, 2004). That is, people are often unaware of what they do not know until they try to explain it, as demonstrated by the illusion of explanatory depth (IOED). First studied by Rozenblit and Keil (2002), the illusion of explanatory depth is the wrongly held belief that one understands aspects of the world on a deeper level than one actually does. This illusion is only revealed to participants when they attempt to generate an explanation for a phenomenon and their true level of understanding is exposed.

The central part of the procedure that Rozenblit and Keil used to elucidate the illusion of explanatory depth is as follows. Participants were presented with a list of devices (e.g., a cylinder lock) and asked to rate their understanding of how each device worked. Next, they were asked to write a detailed, step-by-step explanation of several test items from the list (e.g., explain how a cylinder lock operates) and then again rated their own understanding of that phenomenon. In several different populations, Rozenblit and Keil consistently found that ratings of understanding for how a device worked were significantly lower after generating a detailed explanation than they were initially. This

finding suggests that people do not actually understand the mechanisms of many everyday devices as much as they believe they do and only realize that they lack understanding after attempting to explain.

Importantly, people specifically hold an illusion of their own *understanding* to a much greater extent than they hold an illusion of their own *knowledge*. Consider a computer. A person can have a lot of knowledge about a computer (e.g., knowing the parts and how to use one), without really possessing an understanding of how the computer works. People are better calibrated with their knowledge of facts (e.g., the capital of Brazil), procedures (e.g., how to bake chocolate chip cookies), and narratives (e.g., the plot of *Good Will Hunting*) than their understanding of phenomena with more complex causal relationships (e.g., how a helicopter works; Rozenblit & Keil, 2002). Thus, it is not just that people are generally overconfident about what they know; they specifically have an illusion about their *understanding* of multifaceted causal phenomena. The IOED has been shown in the domains of devices (Lawson, 2006; Mills & Keil, 2004; Rozenblit & Keil, 2002), natural phenomena (Rozenblit & Keil, 2002), and political policy (Alter, Oppenheimer, & Zemla, 2010; Fernbach, Rogers, Fox, & Sloman, 2013).

Explanations for the illusion of explanatory depth

Several explanations have been given for the IOED, which fall broadly into two groups: 1) people overestimate the quality and depth of their mental representations, and 2) people fail to accurately judge their ability to provide good explanations. Within the idea that people overestimate their own mental representations, the IOED has been specifically linked to how people encode complex causal systems. First, explanatory knowledge can have many different levels and people can confuse a surface level

understanding of a causal system for a more detailed one, causing them to hold the illusion that they understand more deeply than they actually do (Keil, 2003; Mills & Keil, 2004; Rozenblit & Keil, 2002). For example, people may know that a cylinder lock works by inserting a key and turning it, which causes the lock to unlock. This understanding might lead people to believe that they know how a lock works, even though they lack an understanding of the more detailed internal mechanisms of the lock. In short, people mistake their shallow representation of the device as a deeper understanding of causal mechanisms.

Additionally, it is much easier for people to explain an entity when it is directly in front of them because they can gain understanding of how an object's mechanisms work by simply examining the object. Often, people have mental representations that they may think are just as real as the actual object. It is only after attempting to generate an explanation that people realize how meager these representations are (Keil, 2003; Mills & Keil, 2004; Rozenblit & Keil, 2002). For example, it is easy to conjure a mental image of a zipper opening and closing but when actually trying to explain a zipper, it becomes clear that this mental image is not actually as helpful as having the object in plain sight. Rozenblit and Keil (2002) found that people held a larger illusion for devices where they perceived a greater proportion of the parts as being visible, suggesting that people may put more faith in their mental images when these images represent devices with more explicitly visible mechanisms.

Not only do people fail to understand the ways in which their mental representations do not accurately reflect reality, they also fail to accurately assess their own ability to provide explanations. First, people are not typically called upon to produce

lengthy, detailed explanations and thus, cannot use their previous experiences to predict their success in generating explanations (Mills & Keil, 2004; Rozenblit & Keil, 2002). People often provide facts and tell narratives, so they can more accurately predict how they will perform on such tasks. However, people may not realize the difficulty of explaining a set of complex mechanisms until they actually attempt to do so, which could lead them to overestimate their ability to generate such explanations.

Additionally, because explanations of complex devices can have many different levels, it can be difficult to judge what levels are necessary to include in a good explanation (Rozenblit & Keil, 2002). When explaining a procedure, one typically knows the end result of the procedure and understands that a good explanation of the procedure would need to include each step needed to achieve this result but would not include something like the mechanics that allow the human body to move to perform the procedure. Contrast this type of explanation with an explanation of a device, e.g., how a flush toilet works. A good explanation might include the way the device is used (e.g., flush the handle and the water is removed), the working of parts inside the tank, the physics behind the water removal, and more. Because it is much more difficult to determine which of these elements are necessary and sufficient for a good explanation, it is hard to easily assess one's ability to generate an explanation without actually doing so. Therefore, people's initial judgments of their ability to explain this type of system are less likely to be accurate.

The IOED and the Division of Cognitive Labor

Studies of the IOED demonstrate the fallacies that people make in evaluating their own understanding. Importantly though, people can and do rely on more than just their

own understanding when generating explanations; they often seek out the understanding of others. People recognize that their own understanding has limits (even if they are inaccurate about where these limits lie) and use their understanding of what others know when choosing who to turn to fill in gaps in their own understanding (Keil, 2012; Wegner, 1987; Wilson & Keil, 1998). Wilson and Keil (1998) first postulated the idea of a division of cognitive labor for explanatory understanding: the idea that any given individual only has a limited understanding of most phenomena and “outsources” the work of understanding these phenomena more deeply to others. These others can be experts in a field, such as when people go to a mechanic with a car problem. However, people often outsource their understanding to non-expert others. For instance, upon facing a car problem, an individual might call a friend who worked on her own car or seek out an Internet forum for people who have had similar car problems. From an early age, children develop skills in navigating the division of cognitive labor by learning to identify which individuals would be able to explain different phenomena (Danovitch & Keil, 2004; Keil et al., 2008; Lutz & Keil, 2002).

The division of cognitive labor is necessary for people to make sense of the world without needing to develop an impossibly overwhelming amount of understanding in their own minds. In fact, it is rare that one’s understanding is evaluated in isolation from others’ understanding, as people can typically ask for advice or search for an answer on the Internet. Recent work suggests that because people so commonly rely on others for understanding, they may actually overestimate the amount of understanding that they possess in seclusion from others. Fisher, Goddu, and Keil (2015) demonstrated that when people used the Internet to search for explanations, they later perceived their ability to

explain a different set of phenomenon as greater compared to people who did not use the Internet. The researchers postulated that searching the Internet led people to conflate an understanding in the world with an understanding in their own mind: that they misjudged where the division of cognitive labor actually lay, believing that they possessed a larger proportion of the total accessible understanding on a topic than they actually did.

Similarly, Kominsky and Keil (2014) showed that people had an illusory belief that they could generate more differences between two similar words than they actually could and that this illusion was larger for pairs where people believed that experts would know more. These findings suggest that the presence of the IOED may be in part facilitated by conflating the understanding of a field that exists outside an individual's own mind with the understanding within one's mind.

Evidence that confusing others understanding and one's own understanding may lead to an illusion can help to explain why the illusion has been shown in previously tested domains. In particular, in the domain of devices, a deep level of understanding exists in the world. For example, there are people who understand devices like cylinder locks very well, even if the average person has a limited understanding of how a cylinder lock works. Thus, the average individual may overestimate their own understanding of a cylinder lock because they recognize that they can outsource their understanding of locks to others. If an illusion arises in part from a confusion of personal and others' understanding, this begs the question of whether people would still demonstrate an illusion in a domain where understanding is limited for everyone. That is, would the illusion exist in a domain where personal understanding and others' understanding should both be low? For instance, in the domain of mental health, even the experts with the

greatest level of understanding currently have fairly little understanding of the basis of disorders and of treatments. In the following, I describe current limits in understanding of the mental health domain. I then explain a series of experiments examining whether perceptions of an overall understanding of a field facilitates an illusion of explanatory depth for that field, focusing particularly on the mental health domain.

Current understanding of mental health

First, I must again reiterate the difference between knowledge and understanding. Many mental health researchers and clinicians know a lot about disorders and treatments but as of yet, this knowledge has not yet translated to a deep or comprehensive understanding. For example, the fundamental way that disorders are defined is still heavily debated. Currently, the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) is the manual used by clinicians and researchers to classify and diagnose mental disorders. The DSM only defines disorders in terms of symptoms but evidence suggests that this approach fails to properly capture real world categories (Barch & Keefe, 2009; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013). Therefore, the DSM model of mental illness has come under criticism as lacking validity from organizations such as the National Institute of Mental Health (Insel, et al., 2010) and the British Psychological Society (2011) and many researchers are pursuing new ways to define disorders more precisely (e.g., Insel & Cuthbert, 2015).

Beyond the lack of consensus about how disorders are defined, current understanding of the causal mechanisms leading to mental health problems is incomplete as well. Disorders are often conceptualized as products of environmental, psychological, and biological factors working together (e.g., Engel, 1977; Kendler, 2005) but little is

truly understood about how these factors interplay to lead to disorders. For instance, researchers can currently identify a number of developmental, genetic and cognitive factors that seem to contribute to the onset and maintenance of schizophrenia (Matheson, Shepherd, & Carr, 2014) but do not yet have a real understanding of how schizophrenia occurs. Research has only scratched the surface of understanding the complex interactions between causal factors of mental disorders.

Even the mechanisms through which the treatments for mental health conditions are effective are unknown. For example, there is a large body of research demonstrating the efficacy of psychotropic drugs, but while it is known that they work by altering brain chemistry, there is still a limited understanding of how these alterations lead to improvements or why medications work differentially across individuals (e.g., Malhotra, Murphy, & Kennedy, 2004; Zhang & Malhotra, 2013). Likewise, the effectiveness of psychotherapy has been consistently demonstrated but the mechanisms that lead to change in therapy are still unknown: research is mixed on whether it is the specific factors of a certain type of therapy (e.g., cognitive-behavioral therapy) or the common factors shared by all therapies (e.g., therapeutic alliance) that make a difference (e.g., Kazdin, 2007; Wampold & Imel, 2015). Again, there is knowledge about the treatments that are effective but not an understanding of why these treatments are helpful.

In the following set of experiments, I examine the IOED in mental health in order to better understand the role that beliefs about others' understanding of a domain may have in creating the IOED. Given that others' understanding may facilitate an illusion of one's own understanding, I examine whether people still show an IOED for mental health and if so, how this illusion compares in magnitude to that previously demonstrated in

devices. To further clarify the possible relationship of such beliefs producing the IOED, I assess whether people's beliefs about others' understanding are related to overestimations of their own understanding.

Overview of Experiments

I conducted four experiments exploring the illusion of explanatory depth in the mental health domain and how that illusion is related to the understanding of others. In Experiment 1, I first examined whether people perceived others' understanding as differing across domains. In Experiment 2, I tested whether these differences in understanding corresponded to differences in the size of the IOED across domains. In Experiment 3, I tested whether explanation was necessary to reveal the illusion in mental disorders to ensure that the demonstrated effect did not result from other factors. Finally, in Experiment 4, I examined whether explicitly manipulating the perceived understanding of others influenced the IOED.

Experiment 1: Establishing Lay Beliefs about Understanding of Mental Health

In order to examine the IOED in mental health as a test of whether others' understanding contributes to an illusion of personal understanding, I first needed to ensure that people actually perceived mental health as less understood by others than devices. Thus, in Experiment 1, I tested beliefs about understanding for mental health items and devices. I tested both disorders and treatments, because both phenomena are poorly understood in the mental health domain. I additionally examined beliefs about medical disorders and treatments, in order to use these domains as comparisons in later tests of the IOED (see Experiment 2 for more information). I examined the perceived societal level of understanding in each domain and predicted that participants would

endorse a lower level of societal understanding for mental items than devices. In addition to examining perceptions of societal understanding, I also examined the perceived gap between lay and expert understanding as another way of assessing beliefs about what others understand. The perceived gap serves as a measure of how clear the division of labor is in a given field: is understanding seen as spread across all others or concentrated within expert others?

Method

Participants. Fifty-one Amazon Mechanical Turk workers participated for pay.

Materials and Procedure. Participants rated five stimuli from each of six different domains: devices, mental disorders, mental disorder treatments, medical disorders, medical disorder treatments, and natural phenomena (see Table 1 for examples of items). Natural phenomena stimuli (e.g., how earthquakes happen) were included in order to increase diversity of items, as in previous IOED work, and will not be discussed further. Mental disorders were selected from independent, previously collected data assessing familiarity and perceived causes to represent disorders that people were familiar with and that people view as psychologically caused, in order to ensure that disorders were seen as actual mental disorders, not as medical conditions.

For each item, participants made two ratings. First, they rated their belief in a societal level of understanding by endorsing their level of agreement with the statement, “We fully understand [ITEM]”, from 1, strongly disagree, to 7, strongly agree (society rating). They also made ratings for the following: “Rate the difference between what an expert (like a college professor or a licensed professional) understands about [ITEM] and what an average person understands about [ITEM]”, from 1, no difference, to 7, very

great difference (gap rating). Finally, participants answered several basic demographic questions. All questions were presented using Qualtrics survey software.

Table 1

Example stimuli in each domain – Experiment 1.

Domain	Item
Devices	How a zipper works
Mental disorders	How the different symptoms of depression develop
Medical disorders	How the different symptoms of a cold develop
Mental treatments	How an antipsychotic works
Medical treatments	How aspirin works

Results

Within each domain, I averaged ratings across all five items in order to create a society score and a gap score for each domain. These scores showed acceptable to good internal consistency across items, all α s > .72. I conducted a 5 (domain: devices vs. mental disorders vs. mental treatments vs. medical disorders vs. medical treatments) repeated-measures ANOVA for each of the two ratings to assess differences across domains. Sidak-corrected follow-up tests were conducted to explore significant effects.

There was a significant effect of domain on society ratings, $F(4, 200) = 72.77, p < .001$ (see Table 2 for means). Follow-up comparisons revealed that devices were seen as more understood than all other domains, all $ps < .001$. Mental disorders were seen as less understood than medical disorders and treatments, $ps < .001$, as were mental treatments, $ps < .035$. Finally, mental disorders were seen as marginally less understood than mental

treatments, $p = .067$. Overall, devices were seen as more understood than medical items, which were in turn seen as more understood than mental items.

There was also a significant effect of domain on gap ratings, $F(4, 200) = 57.53, p < .001$ (see Table 2 for means). People saw the gap between lay and expert understanding as smaller for devices than all other items, all $ps < .001$. Additionally, this gap was seen as marginally larger for mental disorders than for medical treatments, $p = .071$. There were no other domain differences, all $ps > .320$. People perceived the gap between lay and expert understanding as larger for mental items, as well as medical items, than for devices.

Table 2

Mean society/gap ratings by domain – Experiment 1.

	Society rating	Gap rating
Devices	5.72 (0.15)	3.78 (0.16)
Medical Treatments	4.78 (0.17)	5.25 (0.14)
Medical Disorders	4.21 (0.15)	5.43 (0.14)
Mental Treatments	3.75 (0.18)	5.51 (0.16)
Mental Disorders	3.38 (0.16)	5.67 (0.17)

Note: Standard deviations are given in parentheses

Discussion

As predicted, mental disorders and treatments were seen as less understood than devices, which demonstrates that people are at least somewhat aware of the current state of understanding in mental disorders and the fact that it is more limited than in devices. Medical disorders and treatments were seen as more understood than mental health items, but less understood than devices. Additionally, the gap between lay and expert understanding was viewed as larger for mental items than devices, suggesting that an understanding of devices is spread more widely across people than an understanding of

mental health. The gap between lay and expert understanding was equally large for mental and medical items.

These findings justify further testing of the IOED in mental health. The domain differences in perceived societal levels of understanding and lay/expert gaps in understanding mean that examining the IOED across these domains may shed light onto whether the perception that others understand facilitates beliefs about one's own understanding. These results allow us to compare the IOED across domains with various levels of perceived understanding. I further explore this comparison in Experiment 2.

Experiment 2: The IOED in the Mental Health Domain

In Experiment 2, I examined the IOED in mental disorders and mental health treatments using the basic IOED procedure developed by Rozenblit and Keil (2002) and compared the IOED in mental health items (disorders and treatments) to the illusion in devices. I hypothesized that people would show overconfidence in their own understanding because they confuse it with others' understanding. Since people endorse a lower amount of understanding in others for mental health than devices, I predicted that people would have smaller initial ratings for mental health phenomena than for devices. As a result, after explaining, ratings would show a smaller drop for mental items than for devices, demonstrating a smaller illusion than in devices.

In addition to devices, mental disorders, and mental treatments, I also examined the illusion in medical disorders and medical treatments. Given that medical items were between devices and mental items in terms of how well understood by society they were perceived to be, observing the illusion in medical items provides a more nuanced look at exactly how personal and others' understanding are related. It may be that people

demonstrate an illusion in medical items that is smaller than in devices but larger than in mental items. However, the lay/expert gap in understanding is seen as similar for medical and mental items, which may predict a similar sized illusion. Additionally, finding that no illusion exists for either medical or mental disorders could suggest that the illusion simply does not extend to health phenomena.

I also examined the correlations between beliefs about the general level of understanding and ratings of participants' own understanding, both before and after generating an explanation. I expected that if people confuse their own understanding with what others understand, that perceptions of the amount of societal understanding would be correlated with ratings of personal understanding prior to generating an explanation across all domains. However, I did not expect perceptions of societal level understanding to relate to ratings of personal understanding after explaining, because explanations allow people to accurately judge their own understanding, rather than confusing others' understanding with their own.

Additionally, I examined the correlations between beliefs about the gap between lay and expert knowledge and participants ratings of personal understanding. I anticipated that a perception of a smaller gap between lay and expert knowledge would allow for more confusion between personal and others' understanding, given that understanding is more widely spread across individuals, rather than being more concentrated in experts. Thus, I anticipated that perception of a smaller gap would be correlated with higher ratings of personal understanding prior to generating an explanation but not following the generation of an explanation in all domains.

Method

Participants. Participants were 251 Amazon Mechanical Turk workers.

Materials and Procedure. Participants were randomly assigned to see items from one of five domains: devices, mental disorders, mental treatments, medical disorders, or medical treatments. Participants rated five items from the given domain.

Participants learned how to rate their understanding of phenomena using a 1 to 7 scale taken from Rozenblit and Keil (2002). Instructions included examples of a 7, 4, and 1 level understanding of a crossbow. Participants then rated each item on this scale (Time 1 [T1] rating). Then, participants were presented with each item, one by one, and prompted to explain each item in as much detail as possible. Exact instructions are as follows:

“As best you can, please describe all the details you know about [ITEM], going from the first step to the last, and providing the causal connection between the steps. That is, your explanation should state precisely how each step causes the next step in one continuous chain from start to finish. In other words, try to tell as complete a story as you can, with no gaps. Please take your time, as we expect your best explanation.”

After each explanation, participants were asked to again rate their understanding of the item (Time 2 [T2] rating). Finally, participants made several follow-up ratings. For each item, participants made the society and gap ratings as in Experiment 1. However, unlike in Experiment 1, these ratings were completed after participants had assessed their own understanding of items and attempted to explain these items. Additionally, participants answered some basic demographic questions, including questions related to

expertise in the domains of interest. All questions were presented using Qualtrics survey software.

Results

Data analysis. Nine participants had worked in one of these domains in the past (e.g., in a mental health treatment center), so their data were dropped from final analyses, leaving a sample of 242. For T1, T2, society, and gap ratings, I averaged across all five items to create overall mean scores. These scores showed acceptable to good internal consistency across items in all domains, all α s > .69.

The Illusion. In order to examine the illusion across domains, I conducted a 5 (domain: devices vs. medical disorders vs. mental disorders vs. medical treatments vs. mental treatments; between) x 2 (time: T1 vs. T2; within) ANOVA, as in Rozenblit and Keil (2002). I conducted Sidak-corrected follow-up corrections to explore interactions.

Analyses showed a significant effect of time, $F(1, 237) = 62.37, p < .001$, with ratings of personal understanding higher at T1 ($M = 3.31$) than T2 ($M = 2.80$). Thus, averaging across all domains, participants' endorsement of their own understanding decreased after explaining. Results also showed a significant effect of domain, $F(4, 237) = 3.45, p = .009$. However, this effect should be interpreted in light of a significant time X domain interaction, $F(4, 237) = 4.79, p = .001$.

Follow-up comparisons showed the expected illusion in devices, as ratings significantly dropped from T1 to T2, $p < .001$. Additionally, this drop was also significant in mental disorders, $p = .041$, and in mental and medical treatments, $ps < .001$. However, ratings did not drop from T1 to T2 for medical disorders, $p = .669$. Overall,

participants showed an illusion in devices, mental disorders, mental treatments and medical treatments but not medical disorders (see Figure 1).

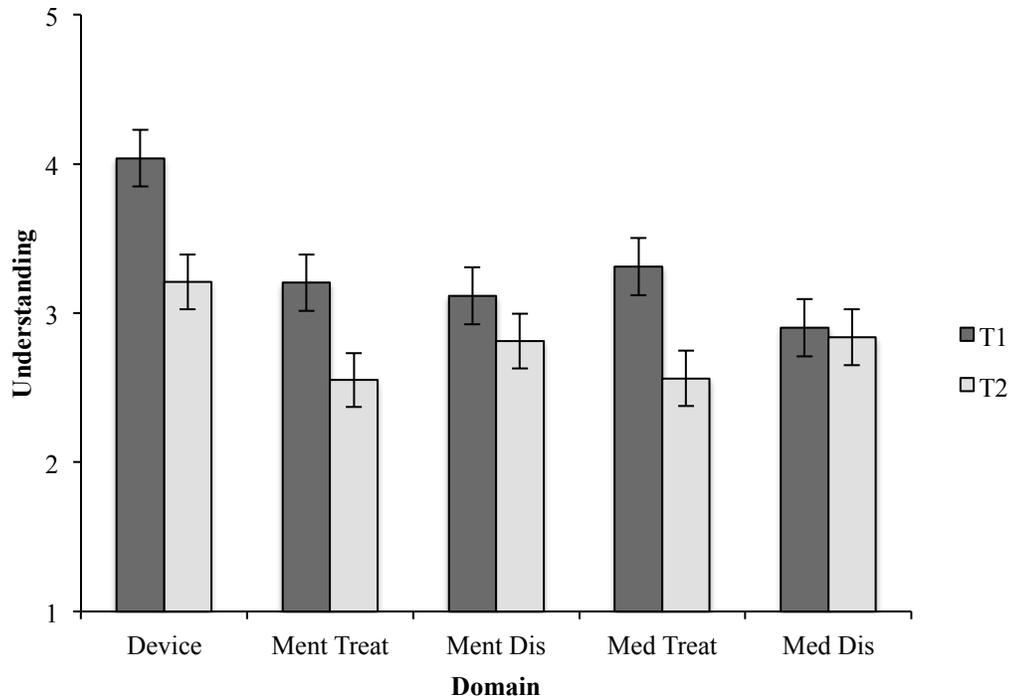


Figure 1. Personal understanding by time – Experiment 2.

In order to compare the size of the illusion across domains, I examined this interaction another way. I predicted that the illusion would be smaller in mental disorders and treatments than in devices. At T1, ratings for devices were marginally higher than those for medical treatments, $p = .077$, and significantly higher than all other domains, all $ps < .022$. However, at T2, ratings did not differ between domains, all $ps > .120$. Given that ratings for devices were greater than other domains at T1 but the same at T2, participants showed a greater drop (and thus a larger illusion) for devices than for mental health phenomena, as expected. The illusion was also smaller in medical treatments than in devices and was non-existent in medical disorders.

Perceived understanding of others. I also examined participants' society and gap ratings (see Table 3). I first compared ratings across domains by conducting a 5 (domain: devices vs. mental disorders vs. mental treatments vs. medical disorders vs. medical treatments; within) ANOVA for both ratings. I predicted that as in Experiment 1, devices would be seen as more understood than medical items, which would in turn be seen as more understood than mental items. There was a significant effect of domain on society ratings, $F(4, 237) = 4.12, p = .003$. Follow-up comparisons showed that devices were seen as more understood than mental treatments, $p = .011$, but were only marginally more understood than mental disorders, $p = .058$, and were not seen as more understood than medical items, $ps > .883$. Mental treatments were seen as less understood than medical disorders, $p = .047$, but not medical treatments, $p = .398$, and mental disorders were not seen as different than medical items, $ps > .188$. Finally, mental disorders and treatments were seen as equally well understood, $p = 1.00$. Unlike in Experiment 1, medical items were not seen as more understood than mental disorders or less understood than devices, and did not differ from one another.

There was also a main effect of domain on gap ratings, $F(4, 237) = 9.96, p < .001$. I predicted that the gap would be seen as smaller for devices than for medical and mental items, as in Experiment 1. Follow-up comparisons showed that people saw the gap between lay and expert understanding as smaller for devices than all other items, all $ps < .002$. There were no differences between other domains, all $ps > .600$. Overall, these results show the same basic findings as Experiment 1, in which devices are seen as more understood and having a smaller lay/expert gap in understanding than mental disorders and treatments.

Table 3

Mean society/gap ratings by domain – Experiment 2.

	Society rating	Gap rating
Devices	4.40 (0.21)	4.85 (0.15)
Medical Treatments	4.27 (0.22)	5.92 (0.16)
Medical Disorders	4.00 (0.22)	6.08 (0.16)
Mental Treatments	3.56 (0.22)	5.71 (0.16)
Mental Disorders	3.41 (0.21)	5.84 (0.15)

Note: Standard deviations are given in parentheses

Next, I examined whether individual's ratings on these two measures correlated with their T1 and T2 ratings. I predicted that society ratings would be positively correlated with T1 ratings but not T2 ratings, and that gap ratings would be negatively correlated with T1 ratings but not T2 ratings. Overall, society ratings were positively correlated with both T1, $r = .221, p = .001$, and T2 ratings, $r = .358, p < .001$, suggesting that higher ratings of other's understanding were related to higher personal understanding at both T1 and T2. Gap ratings were also correlated with both T1, $r = -.249, p < .001$, and T2 ratings, $r = -.214, p < .001$, such that larger perceived differences between lay and expert knowledge were related to lower personal understanding at both T1 and T2. Predictions were supported for correlations at T1 but correlations were also found at T2.

Finally, I examined whether correlations between society/gap ratings and personal understanding held across all domains as expected. For society ratings (see Table 4), higher ratings were correlated with T1 and T2 ratings for devices and mental disorders ($r_s > .280, p_s < .045$). However, higher ratings were only correlated with T2 ratings for medical and mental treatments ($r_s > .325, p_s < .025$) and were only marginally correlated with T2 ratings for medical disorders ($r = .251, p = .089$). Society ratings related more to T2 than T1 ratings across domains and were only correlated with T1 ratings for devices

and mental disorders. Thus, the anticipated relationship between society and T1 ratings was only found in devices and mental disorders.

Table 4

Society rating correlations by domain – Experiment 2.

	T1	T2
Devices	.290*	.412**
Medical Treatments	.150	.326*
Medical Disorders	.131	.251
Mental Treatments	.137	.352*
Mental Disorders	.310*	.400**

* $p < .05$. ** $p < .01$.

For gap ratings (see Table 5), larger perceived gaps were correlated with lower ratings for mental disorders at T1 ($r = -.315, p = .029$) and T2 ($r = -.491, p < .001$).

Larger differences were also marginally correlated with T1 ratings for medical disorders ($r = -.259, p = .078$) and mental treatments ($r = -.271, p = .057$). Differences were not correlated with ratings for devices or medical treatments, $r_s < .036, p_s > .540$. The predicted relationship between gap ratings and personal understanding was found only in the domain of mental disorders, not across all domains.

Table 5

Gap rating correlations by domain – Experiment 2.

	T1	T2
Devices	.035	.011
Medical Treatments	.014	-.090
Medical Disorders	-.259	-.239
Mental Treatments	-.271	.081
Mental Disorders	-.315*	-.491**

* $p < .05$. ** $p < .01$.

Discussion

Experiment 2 provides a first look at people's meta-knowledge in mental health, a scientific domain where understanding is very limited. People show an IOED in mental disorders, mental treatments, and medical treatments, showing that the IOED extends across a number of health domains. Importantly, the illusion was larger in devices than in mental items: ratings were higher than mental items for devices at T1 but not at T2. One potential reason is that people confuse the understanding of the field at large with their own understanding at T1. Because the understanding for mental health is seen as more limited, people initially experience less confusion between others' understanding and their own in mental health domains than in devices. As a result, ratings at T1 are lower for mental items, where people perceive a more limited understanding, and thus, are closer to their own actual levels of understanding for mental disorders and treatments than devices.

People fail to show an illusion for medical disorders, suggesting that people make fairly accurate T1 judgments about their understanding of medical disorders. Medical disorders are seen as less understood than devices and understanding is seen as more concentrated in experts, so one would expect a smaller illusion in medical disorders than devices. In addition, other factors that contribute to the IOED in devices may not be present in medical disorders. People hear explanations for medical disorders much more than for devices, mental items, or medical treatments. Doctors frequently explain to patients how the symptoms of their disorder develop. Because of this, people may have a better idea of what a good explanation looks like for medical disorders, and as a result, may be able to more accurately forecast the quality of explanation that they will be able to generate. In combination with lower levels of others' understanding and a greater

concentration of understanding in experts, a greater accuracy about the contents of a good explanation may help eliminate the illusion in medical disorders. Other possible reasons for the lack of an IOED in medical disorders should be investigated in future research.

In addition to examining the illusion across domains, I also again examined perceptions of others' understanding and the lay/expert gap across domains. While there were still differences in society and gap ratings across domain, there were fewer differences than found in Experiment 1. This is likely due to the fact that society and gap ratings were done after individuals had rated their own understanding and explained, which may have altered these ratings. I have discussed the idea that people's beliefs about others' understanding may influence beliefs about their own understanding but the reverse may also be true: the altered beliefs about personal understanding caused by the experiment may have impacted beliefs about what others understand. In particular, people may have based ratings of others' understanding on their accurate beliefs about their own understanding, explaining the stronger correlations with T2 ratings than T1 ratings.

I also found differences in the relationships between society and gap ratings and personal understanding across domains. I predicted that across all domains, society and gap ratings would be related to personal understanding but find that society ratings are only related to personal understanding at T1 for devices and mental disorders and that gap ratings are only related to personal understanding of mental disorders (at both T1 and T2). Again, one explanation is that in generating an explanation, participants' perceptions of others' understanding were altered, which concealed the true relationships between these constructs. For instance, in the domain of devices, the difficulty of generating an

explanation seems to have lowered society/gap ratings as compared with Experiment 1, which may have occluded the real relationship between prior beliefs about others' understanding and personal understanding.

Experiment 3: Is Explanation Necessary to Reveal the Illusion?

An important part of the IOED phenomenon is the idea that attempting to generate an explanation is necessary to reveal the illusion. Thus, to ensure that the IOED functions the same way in mental disorders as in devices, Experiment 3 examined whether the same drop would occur in personal understanding of mental disorders after simply describing (e.g., listing symptoms, risk factors, etc.) disorders, rather than explaining. A similar paradigm has been deployed in examining people's illusion of understanding for political policies (Fernbach et al., 2013). People were asked to either explain how a policy would work or to list the reasons that they agreed or disagreed with a policy. The IOED was only revealed to those who had explained the policy, not simply listed reasons for their beliefs about a policy. Thus, this experiment serves to extend this basic finding to mental disorders.

Why might simply listing characteristics of mental disorders reveal the illusion for mental health? First, laypeople tend to automatically assume causal relationships between features of phenomena including mental disorders (e.g., Ahn & Kim, 2002). Thus, listing symptoms and factors that could contribute to mental disorders may cause people to spontaneously generate links between these items that could have the same impact of an explanation on their perception of their own understanding. Secondly, people may have more familiarity with the devices that they interact with on a daily basis (e.g., a zipper) than they do with mental disorders. Thus, it may be difficult for people to

even list characteristics of disorders, which could lower their perceptions that they understand how disorders work. Despite these possibilities, I predicted that participants would only show a drop in personal understanding after explaining, not after listing characteristics. In Experiment 3, I compared T1 and T2 ratings when participants were asked to explain mental disorders and simply list characteristics of mental disorders.

Experiment 3 also allowed for a broader investigation of the IOED in mental health. I used a larger set of health disorders in order to examine whether the IOED extends over disorders that people tend to see as biological, in addition to the psychologically caused disorders tested in Experiment 2. Previous work has shown that people tend to view psychological phenomena as less complicated than biological phenomena (Keil, Lockhart, & Schlegel, 2010). Thus, I predicted that people would be more overconfident in their understanding of disorders seen as psychologically caused than disorders seen as biologically caused, and that this will lead to a larger illusion for psychologically caused disorders than biologically caused disorders.

Method

Participants. One hundred Amazon Mechanical Turk workers participated for pay.

Materials and Procedure. Participants were randomly assigned to either the explanation or description condition. Participants in both conditions rated their understanding of eight mental disorders as in Experiment 2. After the initial ratings, participants then either explained or described disorders. The explanation condition was identical to Experiment 2, with the only change being that participants rated eight disorders, rather than five. Participants in the description condition were asked to list all

of the characteristics that they know about each disorder, rather than explaining each disorder. Exact instructions are as follows:

“As best you can, please list all the characteristics you know about [ITEM] including things like symptoms, typical attributes of someone with the disorder, risk factors and so on. Try to list as many characteristics as you can. Press the return key between each separate item that you list. Please take your time, as we expect your best list.”

After listing characteristics, participants again rated their understanding (T2 rating). Then, all participants completed several follow-up ratings. For each disorder, they completed three ratings assessing how biologically, psychologically and environmentally caused the disorder was, on a scale from 0, not at all biologically/psychologically/environmentally caused, to 100, completely biologically/psychologically/environmentally caused. Additionally, participants completed society, gap and demographic questions as in Experiment 2.

Results

Data analysis. Two participants endorsed working in a mental health setting, so their data was dropped from final analyses, leaving a sample of ninety-eight. For T1, T2, society, and gap ratings, I averaged across all eight items to create overall scores.

The illusion. To examine whether description can reveal the illusion in mental disorders, I conducted a 2 (condition: understanding vs. knowledge) x 2 (time: T1 vs. T2) mixed ANOVA. Results showed a significant effect of time, $F(1, 96) = 6.84, p = .010$, with ratings higher at T1 ($M = 2.86, SD = 0.13$) than T2 ($M = 2.63, SD = 0.13$). There was also a significant effect of condition, $F(1,96) = 4.21, p = .043$, with personal

understanding rated lower in the explanation condition ($M = 2.49$, $SD = 0.18$) than the description condition ($M = 3.00$, $SD = 0.17$).

These effects should be interpreted in light of a significant interaction between time and condition, $F(1, 96) = 11.71$, $p = .001$. I predicted that ratings would drop from T1 to T2 in the explanation condition but not in the description condition. Follow-up Sidak-corrected comparisons show that ratings significantly dropped from T1 to T2 in the explanation condition, $p < .001$, but not in the description condition, $p = .564$ (see Figure 2). Examining this interaction another way, ratings between conditions did not differ at T1, $p = .454$, but are significantly lower at T2 for the explanation condition than the description condition, $p = .002$. Thus, participants in both conditions endorsed the same initial level of understanding but only those who explained their understanding showed a drop at T2. The illusion was only demonstrated in participants who generated explanations for mental disorders, not those who simply listed characteristics.

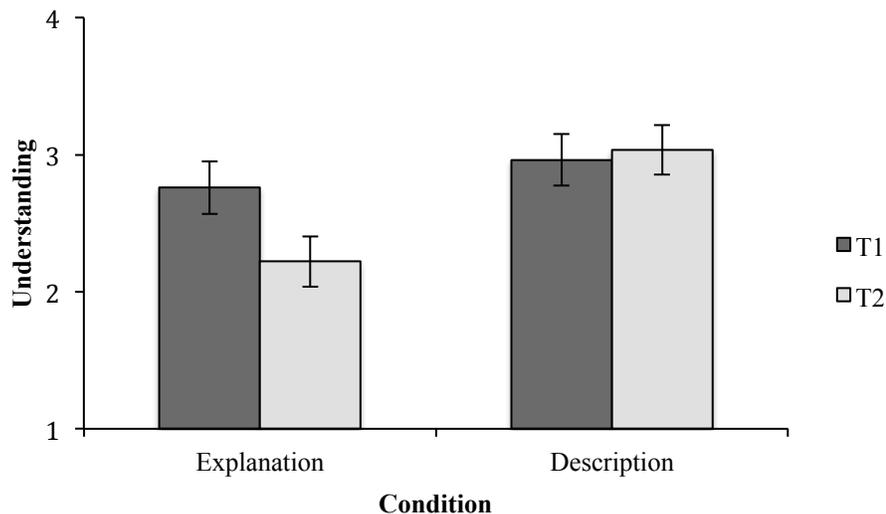


Figure 2. Personal understanding across condition – Experiment 3.

Perceived basis of disorder. I also examined the effect of the perceived basis of the disorder on T1 and T2 ratings. To do so, I first classified the four disorders with the highest scores on the biologically caused question and the lowest scores on the psychologically caused question (ADHD, schizophrenia, bipolar disorder, borderline personality disorder) as biological and the four disorders with the highest scores on the psychologically caused question and the lowest scores on the biologically caused question (generalized anxiety disorder, OCD, depression, anorexia) as psychological. Then, I calculated T1 and T2 ratings for biological and psychological disorders. Because I was interested in how causal factors were related to the illusion, I only examined the explanation condition. I conducted a 2 (cause: biological vs. psychological) x 2 (time: T1 vs. T2) mixed ANOVA.

Results showed the expected significant effect of time, $F(1, 47) = 14.40, p < .001$, with ratings lower at T2 than T1. There was also a significant effect of cause, $F(1, 47) = 56.04, p < .001$, with ratings higher for psychological than biological disorders. Finally, there was significant cause by time interaction, $F(1, 47) = 5.04, p = .029$. For biological disorders, the drop from T1 to T2 was significant, $p = .008$, but smaller than the drop for psychological disorders, $p < .001$ (see Figure 3). Thus, as predicted, people demonstrate a larger illusion in psychologically caused disorders than in biologically caused disorders.

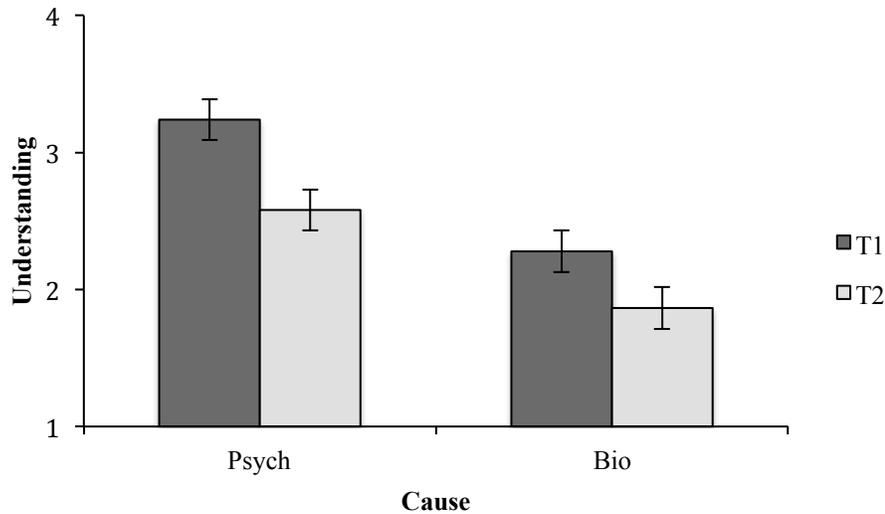


Figure 3. Personal understanding by disorder cause – Experiment 3.

Perceived understanding of others. First, I compared society and gap ratings across conditions. Neither society ratings $M = 3.70$, $SD = 1.32$, $t(96) = .037$, $p = .759$, nor gap ratings, $M = 5.74$, $SD = 1.13$, $t(96) = .996$, $p = .648$ differed across conditions. I then examined correlations split by condition. For the description condition, society ratings were correlated with T1, $r = .370$, $p = .008$, and T2 ratings, $r = .408$, $p = .003$, and gap ratings were uncorrelated with T1 ratings, $r = -.171$, $p = .234$, and marginally correlated with T2 ratings, $r = -.269$, $p = .059$. In the explanation condition, participants showed a different pattern. Society ratings were uncorrelated with T1 or T2 ratings, $r_s > .190$, $p_s > .135$, and gap ratings were correlated with T1, $r = -.294$, $p = .042$, and T2 ratings, $r = -.316$, $p = .029$. Higher ratings of societal understanding were related to higher ratings of personal understanding but only in the description condition. Likewise, ratings of a larger lay/expert gap were negatively related to personal understanding only in the explanation condition.

Discussion

The major finding from this experiment is that, as predicted, explanation seems to be truly necessary to reveal the illusion. Simply thinking about the various aspects of a phenomenon is not enough to reveal one's true level of understanding: those who did so remained just as overconfident about their understanding. People did not seem to spontaneously consider the causal links between the symptoms and risk factors that they generated, so creating a list was not useful in revealing one's true level of understanding. This finding further demonstrates the useful nature of explanations in helping people to evaluate their true level of understanding in a way that other types of thinking do not. The illusion of explanatory depth seems to function the same way in mental disorders as it does in previously studied phenomena.

Additionally, this experiment demonstrates that the IOED extends to a wider range of disorders than previously tested, even disorders that vary in perceived causes. However, ratings of understanding do differ depending on the perceived cause of the disorder, with biological disorders being less understood than psychological disorders, both before and after explanation, leading to a smaller illusion in biological disorders, as predicted. People tend to believe that psychological concepts are easier to learn through experience than are biological concepts (Keil et al., 2010). As such, people may believe that they have greater insight into disorders that they see as psychological because they have more experience with psychological concepts than biological. Thus, people may be more prone to overestimate their own understanding of psychological disorders.

Interestingly, explaining disorders seemed to actually change people's perceptions of what causes individual mental disorders in comparison to previous research. In prior work by Ahn, Proctor, and Flanagan (2009), people tended to rate several of the disorders

that I used as either extremely biological (e.g., schizophrenia) or extremely psychological (e.g., anorexia). But in this experiment, people moderately endorsed both psychological and biological causes for all disorders. Importantly, in Experiment 3, people only rated the causal basis of disorders after they had already attempted to explain how the symptoms of these disorders had developed. It is possible that being asked to explain a mental disorder may cause people to reflect on the wide variety of causal factors that could play a role, leading them to rely less on heuristics that class disorders as either biological or psychological. However, comparing my results and the previous research of Ahn et al. is a cross-experimental comparison, and so is only suggestive of possible differences. Future research should examine whether attempting to explain a mental disorders causes people to instantiate a wider variety of causes for that disorder.

Finally, I found that the relationships between perceived understanding of others and individual's own understanding ratings differed across conditions. Personal understanding at both T1 and T2 was correlated with ratings of societal understanding only in the description condition and with gap ratings only in the explanation condition. This finding provides further evidence that the experiment actually impacted people's ratings of others' understanding. In particular, the fact that correlations differed across conditions for T1 ratings demonstrates that the act of either explaining or describing impacted later society/gap ratings. The conditions were only different after T1 ratings were made: thus, relationships should have been the same across conditions if society/gap ratings had not been altered. Despite the difficulty of interpreting these results, overall, it seems that ratings of societal understanding and the lay/expert gap are related to individual's ratings of understanding.

Experiment 4: Manipulating Perceived Scientific Understanding

So far, the proposed experiments have tested the illusion in mental health as related to the role that a field's perceived understanding of a phenomenon plays in people's perception of their own understanding of that phenomenon. In Experiment 4, I directly manipulated whether others are described as having high level or a low level of understanding of mental disorders in order to more directly test whether perceptions that others understand lead to an illusion of personal understanding. In light of the results of Experiment 2, which showed that people had a larger illusion for the most well understood field, I predicted that learning that mental disorders are well understood would lead people to overestimate their own personal understanding initially. Higher initial ratings would lead to a larger drop after explaining and a larger illusion. In contrast, I predicted that learning that there is a low level of understanding of mental disorders would reduce or eliminate the illusion: people would be unable to confuse others' understanding with their own, since both are limited.

Method

Participants. Participants were 101 Amazon Mechanical Turk workers participated for pay.

Materials and Procedure. Prior to completing ratings of understanding as in the previous experiments, participants were randomly assigned to read that scientists had a high level of understanding (high condition) or a low level of understanding (low condition) of mental health. Participants first read an artificial description of the current scientific understanding of a generally unfamiliar disorder (i.e., cyclothymic disorder) and were told that this understanding reflected the level of scientific understanding for

other mental disorders. In the high condition, scientists were portrayed as understanding the disorder well, while in the low condition, scientists were portrayed as lacking understanding (see Appendix A for exact descriptions).

Results

Data analysis. Two participants endorsed working in a mental health setting, so their data was dropped from final analyses, leaving a sample of ninety-nine participants. For T1, T2, society, and gap ratings, I averaged across all eight items to create overall scores.

Quality of manipulation. First, I examined ratings related to the manipulation itself (see Table 6). Our manipulation check questions showed that the manipulation was effective in changing agreement that scientists understood cyclothymic disorder, $t(93.66) = 8.33, p < .001$, and other mental disorders, $t(93.22) = 4.55, p < .001$, with ratings higher in the high than the low condition. Additionally, the described understanding was seen as equally likely to extend to other mental disorders, $t(97) = 1.43, p = .156$, and to apply to the same percentage of other disorders, $t(95) = 1.05, p = .298$, across condition. As expected, the high description was seen as a better quality explanation than the low description, $t(97) = 3.50, p = .001$. The high description was also seen as more plausible than the low description, $t(96) = 2.81, p = .006$.

Table 6

Mean manipulation ratings by condition – Experiment 4.

	Society rating	Gap rating
Understand Cyclothymic	5.02 (1.22)	2.78 (1.45)
Understand Other Ment	4.56 (1.28)	3.27 (1.54)
Extend to Other Ment	61.08 (24.04)	54.12 (24.35)
Percentage of Other Ment	52.53 (22.53)	47.73 (22.66)
Quality of Explanation	4.46 (1.74)	3.33 (1.46)
Plausibility	70.00 (24.49)	56.52 (22.91)

Note: Standard deviations are given in parentheses

The illusion. In order to compare the illusion across high and low understanding conditions, I conducted a 2 (condition: high or low understanding; between) x 2 (time; within) mixed ANOVA. Results showed a significant effect of time, $F(1, 97) = 16.83, p < .001$, with ratings higher at T1 ($M = 3.15$) than T2 ($M = 2.65$) but no effect of condition, $F(1,97) = 1.02, p = .304$. However, results showed a significant interaction, $F(1, 97) = 6.05, p = .016$. Follow-up comparisons showed a drop from T1 to T2 in the high condition, $p < .001$, but not in the low condition, $p = .251$ (see Figure 4). Examining the interaction in another way, ratings were higher in the high condition than the low condition at T1, $p = .046$, but not at T2, $p = .885$. As anticipated, participants showed an illusion when others were described as having a high level of understanding but not a low level of understanding.

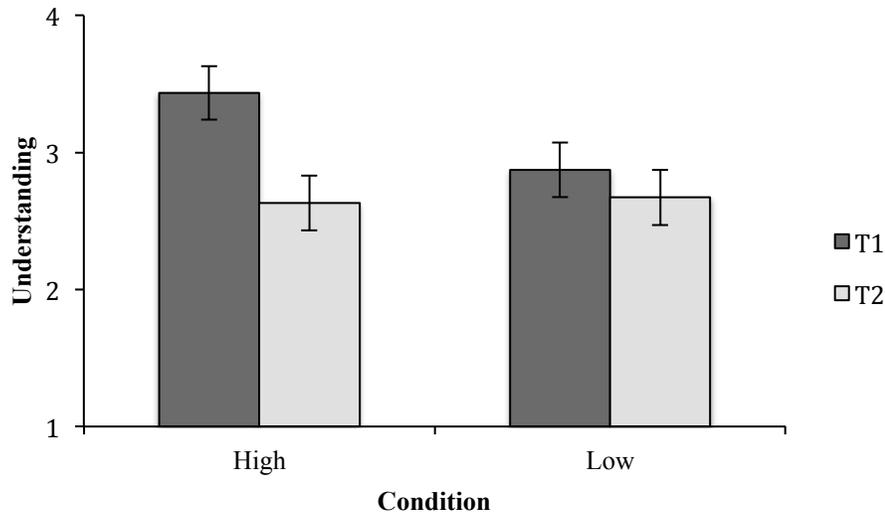


Figure 4. Personal understanding across condition – Experiment 4.

Perceived understanding of others. I compared society and gap ratings across conditions. Neither society ratings, $M = 3.59$, $SD = 1.34$, $t(97) = 1.34$, $p = .183$, nor gap ratings, $M = 5.70$, $SD = 1.05$, $t(97) = 0.31$, $p = .755$, differed across conditions. Then, I examined the correlations between these ratings, T1 ratings and T2 ratings, split by condition. In the high condition, society ratings were significantly correlated with T1 ($r = .334$, $p = .018$) and T2 ratings ($r = .387$, $p = .006$), and gap ratings were significantly correlated with T2 ratings ($r = -.450$, $p = .001$). In the low condition, society and gap ratings were not correlated with T1 or T2 ratings (all $ps > .165$). Thus, ratings of societal level understanding and the lay/expert gap were only related to ratings in the high condition, not the low condition.

Discussion

Experiment 4 shows that directly manipulating the perceived understanding of others influences people’s perceptions of their own understanding. In particular, learning that a field is well understood seems to inflate initial ratings of understanding, leading to

a larger drop in understanding after explaining. On the other hand, the illusion is actually eliminated when people learn that a field is not well understood. This may be because people are then unable to confuse what they know with what others know, so they initially judge their own understanding more accurately. Along with results of Experiment 2, this finding provides converging evidence that the illusion is in part due to confusion between others' understanding and one's own personal understanding.

Ratings about the manipulation showed that people believed that the described understanding of cyclothymia was equally likely to extend to other disorders and to the same percentage of disorders, regardless of condition. This shows that participants were not simply applying information from the description to more disorders in one condition than the other: the information was equally relevant to other disorders across condition. One issue with the descriptions was that they were unbalanced in terms of plausibility, with the well understood description seen as more plausible than the low understanding description. A useful future study would match descriptions on plausibility, to ensure that it is truly the level of understanding described that impacts people's ratings.

Ratings of others' understanding and perceived lay/expert gaps in understanding were only related to self-reports of understanding in the high condition, not in the low condition. This finding helps to provide further evidence that believing that others understand is an important component of the IOED. When others were portrayed as having a larger understanding of mental disorders, people may have confused others understanding with their own understanding, leading to relationships between ratings of other and self understanding. However, in the low condition, people may have based

ratings of their own understanding on factors aside from others' understanding, because others' understanding was lacking.

General Discussion

The described experiments examined the illusion of explanatory depth in mental health, a domain where an understanding of complex causal mechanisms is extremely limited. Experiment 1 demonstrated that laypeople do indeed see mental disorders and treatments as less well understood generally than devices. Experiment 2 showed that people held an IOED for both mental disorders and mental health treatments, and that these illusions were smaller than in the device domain. Experiment 3 revealed that explanation is necessary to illuminate the illusion in mental health and that just listing characteristics is not enough. Finally, Experiment 4 provided further evidence that believing that others understand underlies the illusion.

I have provided evidence that perceptions of others' understanding are an important cornerstone of the IOED. As discussed, a potential mechanism for the IOED is that people confuse what others understand with what they understand on their own and thus, overestimate their own understanding (e.g., Fisher et al., 2015). As shown by the findings of Experiment 2, in mental health, where others' understanding is perceived as lower, people demonstrate a smaller illusion than in devices. People believe that there is less understanding overall, so even if they mistakenly believe that they possess a larger portion of this understanding than they actually do, they still believe that they understand less than they do in devices. Additionally, describing a lack of understanding in the field of mental health eliminated the illusion (Experiment 4). If people believe that the understanding of others is extremely limited, they are unable to misattribute a greater

portion of others' understanding to themselves and are thus more accurate about their own understanding.

One remaining question is why people did not show an illusion for medical disorders. As discussed earlier, in addition to being seen as less understood than devices, people may be more likely to experience learning a good explanation for a medical disorder, which could make them more accurate about their own ability to explain (Mills & Keil, 2004; Rozenblit & Keil, 2002). The finding from Experiment 3 that people show less of an illusion for mental disorders that they see as biologically caused may also shed light onto the lack of illusion in medical disorders. As noted, biological phenomena are seen as more difficult to understand by laypeople than are psychological phenomena (Keil et al., 2010). If something is seen as easier to understand generally, people may be more likely to confuse their understanding with that of others. Thus, in medical disorders, there may be several factors limiting the presence of an illusion.

In extending the IOED to mental disorders, I also found that it functions similarly to the illusion in devices, in that only explanations are able to reveal the illusion (Experiment 3). Importantly, past work has mainly assessed the importance of explanation by showing that people show an illusion in domains with more complicated explanations (e.g., devices) but not in those where explanations are easier to generate (e.g., movie narratives; Rozenblit & Keil, 2002). This is one of the only experiments that directly compares explanation to another type of higher-order thinking (listing characteristics) within the same domain. Additionally, this finding advances the study of the IOED by supporting causal relationships as one of the fundamental aspects of explanations. People who were asked to list disorder characteristics generated symptoms

and risk factors, which would be important in a good explanation of a disorder. However, they did not elucidate the causal relationships between these factors, which seems to have eliminated the power of explanation to illuminate the illusion.

Beyond examining the relationship between others' understanding and the IOED, I also demonstrate that people hold an illusion for a wider range of domains than previously tested, including mental disorders and treatments. This domain extension demonstrates the robustness of the IOED phenomenon. People do not seem to overestimate their understanding in just a few select domains; instead, they consistently hold illusions of their own understanding across a number of disparate phenomena. In particular, the finding that people hold an IOED for mental disorders shows that people not only demonstrate an illusion for external, non-human phenomena (e.g., devices) but also for phenomena that happen within people. This extension across a larger number of domains is a novel one that supports the fundamental nature of the IOED.

One particularly interesting implication of the finding that people have an IOED in the mental health domain is that mental disorders are internal and invisible to observers. In the past, visibility has been linked to the IOED, with people demonstrating a larger illusion for devices with more visible parts (Rozenblit & Keil, 2002). Additionally, one factor described as underlying the IOED is that people wrongly believe that their mental images of a device are just as useful in understanding as actually seeing the device. Yet I find that even in mental disorders, where processes are not at all visible, people still hold an illusion. This finding suggests that while perceptions of visibility may contribute to the presence of an illusion, they are not necessary for people to hold an IOED.

Implications

This set of experiments has several implications, for clinical health and expert communication. First, by confirming that an IOED exists in mental health and demonstrating that explanation is an important tool in achieving accuracy about one's own understanding of mental health, these findings suggest that explanations should be used by both clinicians and patients in dealing with mental health issues. For example, clinicians frequently use data from the client to produce a conceptualization of the client's problems, which is then used to guide the treatment process (e.g., Sperry, 2005; Sperry & Sperry, 2012). It is important that in these conceptualizations, the clinician fully elaborates the causal relationships between the client's problems, as well as other important factors in the client's life, in order to ensure that they have a complete understanding of the case. The same is true for clients: as a central part of therapy, clients often focus on generating an explanation for why they feel, think and act as they do (Guignon, 1998). Findings from Experiment 3 suggest that it is important that these explanations are explicit and coherent: if they simply represent descriptions of problems, without causal connections, clients may be overconfident in their own understanding, leading them to overlook aspects of their problems for which they do not have a good understanding.

Additionally, evidence that an illusion of understanding may be facilitated by a perception that others understand has implications for how experts communicate their level of understanding. In particular, it suggests that overstating the current state of the field's understanding may not only be deceptive, but could also serve to make people overconfident in their own understanding. This could have a particularly detrimental

impact in the field of mental health. In the United States, one in five people is diagnosable with a mental illness yearly (SAMHSA, 2014) and an overconfidence in one's understanding may reduce one's impetus to seek treatment or information related to mental health. More broadly, experts should recognize that the way that they communicate their understanding may have large impacts on how people see their own understanding.

Future Directions

My experiments open several avenues for future research. First, studying the illusion in mental disorders opens up the possibility for examining the IOED in experts in mental health. Because even experts only have a limited understanding, it is interesting to consider whether experts, like clinicians, overestimate their own understanding of mental disorders. Potentially, the same mechanisms may play a role: if clinicians perceive the field as a whole as being well understood, they may be more likely to confuse this understanding with their own and to show an IOED. On the other hand, clinicians may be able to accurately assess their own understanding, without showing an illusion. The IOED has not yet been studied in experts and this would be a valuable addition to the literature.

My experiments suggest that people's perceptions of others' understanding impact how they see their own understanding. However, what is less clear is how people actually make their judgments about what others understand. Previous work has examined how people decide who has expertise in a domain (Danovitch & Keil, 2004; Keil et al., 2008; Lutz & Keil, 2002) but not how they assess the general level of understanding for a given domain. In particular, my findings suggest that assessing one's

own level of understanding can actually alter beliefs about levels of societal understanding, suggesting a potential reciprocal relationship between perceptions of one's own and of others' understanding. Further work should the nature of this relationship and other factors that determine perceptions of others' understanding.

Generally, this set of experiments further extends work demonstrating the robustness of the IOED phenomenon and the phenomenon's extent over multiple domains, including mental disorders, medical treatments, and mental health treatments. Additionally, the experiments provide support for a mechanism that has only recently been examined: the idea that people mistake others' understanding for their own and that this is what leads people to initially overestimate their understanding. As research on mental health issues progresses, understanding will only increase but until an infinitely distant future, people must work with the limited understanding that they possess: this set of experiments helps to better explain how people do so.

References

- Kim, N. S., & Ahn, W. K. (2002). The influence of naive causal theories on lay concepts of mental illness. *American Journal of Psychology*, *115*, 33-66.
- Ahn, W. K., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science*, *33*, 147-182.
- Alter, A. L., Oppenheimer, D. M., & Zemla, J. C. (2010). Missing the trees for the forest: A construal level account of the illusion of explanatory depth. *Journal of Personality and Social Psychology*, *99*, 436-451.
- Barch, D. M., & Keefe, R. S. (2009). Anticipating DSM-V: Opportunities and challenges for cognition and psychosis. *Schizophrenia Bulletin*, *36*, 43-47.
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *36*, 421-441.
- British Psychological Society. (2011). Response to the American Psychiatric Association: DSM-5 development. Retrieved from http://apps.bps.org.uk/_publicationfiles/consultation-responses/DSM-5%202011%20-%20BPS%20response.pdf
- Danovitch, J. H., & Keil, F. C. (2004). Should you ask a fisherman or a biologist?: Developmental shifts in ways of clustering knowledge. *Child Development*, *75*, 918-931.
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*, 83-87.

- Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, 99, 3-19.
- Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, 196, 129-136.
- Fernbach, P. M., Rogers, T., Fox, C. R., & Sloman, S. A. (2013). Political extremism is supported by an illusion of understanding. *Psychological Science*, 24, 939-946.
- Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for explanations: How the Internet inflates estimates of internal knowledge. *Journal of Experimental Psychology: General*, 144, 674-687.
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69, S342-S353.
- Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system. In F. Keil & R. Wilson (Eds.). *Cognition and explanation*, (pp. 299-323). Cambridge, MA: MIT Press.
- Guignon, C. (1998). Narrative explanation in psychotherapy. *American Behavioral Scientist*, 41, 558-577.
- Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15, 135-175.
- Insel, T. R., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science*, 348, 499-500.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167, 748-751.

- Iselin, M. G., & Addis, M. E. (2003). Effects of etiology on perceived helpfulness of treatments for depression. *Cognitive Therapy and Research, 27*, 205-222.
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27.
- Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology, 57*, 227-254.
- Keil, F. C. (2012). Running on empty? How folk science gets by with less. *Current Directions in Psychological Science, 21*, 329-334.
- Keil, F. C., Lockhart, K. L., & Schlegel, E. (2010). A bump on a bump? Emerging intuitions concerning the relative difficulty of the sciences. *Journal of Experimental Psychology: General, 139*, 1-15.
- Keil, F.C., Rozenblit, L.R. and Mills, C. (2004). What lies beneath? Understanding the limits of understanding. In D.T. Levin (Ed.), *Thinking and Seeing: Visual Metacognition in Adults and Children*, (pp. 227-249). Cambridge: MIT Press.
- Keil, F. C., Stein, C., Webb, L., Billings, V. D., & Rozenblit, L. (2008). Discerning the division of cognitive labor: An emerging understanding of how knowledge is clustered in other minds. *Cognitive Science, 32*, 259-300.
- Kendler, K. S. (2005). Toward a philosophical structure for psychiatry. *American Journal of Psychiatry, 162*, 433-440.
- Kim, N. S., & Ahn, W. K.. (2002). The influence of naive causal theories on lay concepts of mental illness. *American Journal of Psychology, 115*, 33-66.
- Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48*, 507-531.

- Kominsky, J. F., & Keil, F. C. (2014). Overestimation of knowledge about word meanings: The “misplaced meaning” effect. *Cognitive Science*, *38*, 1604-1633.
- Kosslyn, S. M. (1995). Freud returns? In R. L. Solso & D. W. Massaro (Eds.), *The science of the mind: 2001 and beyond* (pp. 90-106). New York: Oxford University Press.
- Lawson, R. (2006). The science of cycology: Failures to understand how everyday objects work. *Memory & Cognition*, *34*, 1667-1675.
- Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, *10*, 464-470.
- Lutz, D. J., & Keil, F. C. (2002). Early understanding of the division of cognitive labor. *Child Development*, *73*, 1073-1084.
- Malhotra, A. K., Murphy Jr, G. M., & Kennedy, J. L. (2004). Pharmacogenetics of psychotropic drug response. *American Journal of Psychiatry*, *161*, 780-796.
- Marsh, J. K., & Romano, A. L. (2015). Lay judgments of mental health treatment options: The mind vs. body problem. Manuscript submitted for publication.
- Matheson, S. L., Shepherd, A. M., & Carr, V. J. (2014). How much do we know about schizophrenia and how well do we know it? Evidence from the Schizophrenia Library. *Psychological Medicine*, *44*, 3387-3405.
- Mills, C. M., & Keil, F. C. (2004). Knowing the limits of one’s understanding: The development of an awareness of an illusion of explanatory depth. *Journal of Experimental Child Psychology*, *87*, 1-32.
- Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*, 521-562.

- Salmon, W. C. (1989). Four decades of scientific explanation. *Scientific Explanation*, 13, 3-219.
- Sperry, L. (2005). Case conceptualization: A strategy for incorporating individual, couple and family dynamics in the treatment process. *The American Journal of Family Therapy*, 33, 353-364.
- Sperry, L., & Sperry, J. J. (2012). *Case conceptualization: Mastering this competency with ease and confidence*. New York: Routledge.
- Strevens, M. (2008). *Depth: An account of scientific explanation*. Cambridge, MA: Harvard University Press.
- Substance Abuse and Mental Health Services Administration (2014). *Results from the 2013 national survey on drug use and health: Mental health findings*, (NSDUH Series H-49, HHS Publication No. (SMA) 14-4887). Rockville, MD: Substance Abuse and Mental Health Services Administration.
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: Research evidence for what works in psychotherapy* (2nd ed.). New York: Routledge.
- Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York, NY: Springer-Verlag.
- Wilson, R. A., & Keil, F. (1998). The shadows and shallows of explanation. *Minds and Machines*, 8, 137-159.
- Zhang, J. P., & Malhotra, A. K. (2013). Pharmacogenetics of antipsychotics: Recent progress and methodological issues. *Expert Opinion on Drug Metabolism & Toxicology*, 9, 183-191.

Appendix A: Experiment 4 Understanding Descriptions

High Understanding Description

“Scientists know a lot about how many mental health disorders work. For instance, they now know a lot about the way that cyclothymic disorder works. They have located several genetic mutations that can cause this disorder and determined that these genetic mutations are passed down through families. However, they also know that an individual with these genes will only experience symptoms of cyclothymic disorder if they experience a certain threshold of negative events. Scientists are sure that these events typically involve trauma but if someone is born with more genetic mutations, more minor stressors can lead them to develop symptoms. When someone with genetic mutations experiences these negative stressors, their brain chemistry changes in response, and scientists are sure that this is what leads to symptoms. Additionally, sleeping patterns can influence how strong symptoms are – individuals who are chronically sleep-deprived display worse symptoms. Overall, scientists have a fairly good understanding of this disorder. Researchers of other mental disorders have reached similar levels of understanding for those disorders.”

Low Understanding Description

“Scientists still do not know a lot about how many mental health disorders work. For instance, they still do not know a lot about the way that cyclothymic disorder works. They believe that several genetic mutations can cause this disorder but do not know what these genetic mutations are or where they come from. However, it also seems that the symptoms of cyclothymic disorder are more likely to occur if someone experiences certain negative events. Scientists are not sure how strong these negative experiences must be for a person to develop symptoms – they might be traumatic events or more minor stressors. Somehow, genetics and negative events seem to change the brain chemistry of people with cyclothymic disorder, although scientists are not sure whether this is what leads to symptoms. Additionally, sleeping patterns seem to be related to symptoms but the relationship is unclear. Overall, scientists do not have a good understanding of this disorder. Researchers of other mental disorders have similar levels of lack of understanding for those disorders.”

Andrew S. Zeveney
811 Meade Street, Bethlehem, PA 18015
302-632-4541 • azeveney@gmail.com

EDUCATION

Lehigh University, Bethlehem, PA

M.S. in Psychology, May 2016

GPA: 3.91

Master's Thesis: Illusion of Understanding in a Misunderstood Field: Lay Beliefs about their Understanding of Mental Health

Select courses: Analysis of Experimental Data, Longitudinal Data Analysis

B.S. in Psychology, May 2014

GPA: 3.81

Honors Thesis: The Impact of Multiple Causes on Categorization

RESEARCH EXPERIENCE

Graduate Student Researcher, Lehigh University 8/15 – present

Categorization and Causal Reasoning Lab – Department of Psychology

- Designing experiments and administering them through undergraduate research pool and Amazon's Mechanical Turk.
- Constructing datasheets using Excel and analyzing data using SPSS.
- Managing undergraduate research assistants in stimuli development and coding.

Researcher, Lehigh University 7/15 – present

Work Hard, Play Smart Program

- Assisted with the development of the content of a novel intervention program to reduce negative consequences of drinking in first-year students.
- Developed measure to assess facilitator fidelity to intervention content.
- Observing and coding intervention sessions for fidelity.

Lab Manager, Lehigh University 1/15 – 8/15

Categorization and Causal Reasoning Lab – Department of Psychology

- Supervised daily activities of undergraduate research assistants.
- Wrote and managed IRB submissions for ongoing studies.
- Recruited physicians to participate in ongoing studies by contacting physician's offices and hospitals.

Research Assistant, Lehigh University 1/15 – 5/15

Universal Screening Lab – Department of Education

- Assembled survey packets to distribute to students in local schools.
- Conducted preliminary data analyses using SPSS.

Researcher, Lehigh University 5/13 – 8/13; 5/14 – 12/14

Categorization and Causal Reasoning Lab – Department of Psychology

- Executed grant-funded project examining lay beliefs about the nature of mental disorder by designing initial studies.
- Coded data and managed data using Excel and analyzed data with SPSS.
- Developed and pretested stimuli for use in future studies.

Research Assistant, Lehigh University 8/12 – 5/13; 8/13 – 5/14
 Categorization and Causal Reasoning Lab – Department of Psychology

- Designed and programmed studies using Qualtrics survey software.
- Analyzed data using Excel and SPSS.
- Conducted literature searches and coded free response data.

Research Assistant, Lehigh University 5/13 – 8/13
 Group Processes Lab – Department of Psychology

- Recruited on-campus participants for study examining perceptions of racial groups.
- Lead participants through procedures, using E-Prime recording software.
- Managed reaction time data and coded for errors.

AWARDS AND HONORS

Lehigh University Presidential Scholarship, 2014

- full year tuition scholarship awarded for academic excellence

Williams Prize in Psychology, Sociology and Anthropology, 2014

- awarded for excellence in research writing

Dean's List, 2011-2014

R.C. Eckardt Scholar, 2010-2014

- honors program designed for high-achieving students to participate in independent work

PUBLICATIONS

Zeveney, A., & Marsh, J. K. (under review). The Impact of Multiple Routes to Category Membership on Essentialism.

Marsh, J. K., & Zeveney, A. (2015). Naïve beliefs about intervening on causes and symptoms in

the health domain. In Noelle, D. C., Dale, R., Warlaumont, A. S., Yoshimi, J., Matlock, T., Jennings, C. D., & Maglio, P. P. (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society

SCHOLARLY PRESENTATIONS

Zeveney, A., & Marsh, J. K., (2016, May). Lay Use of Informant Discrepancies in Evaluating

Mental Health Treatments. In A. De Los Reyes (Chair), *Low correspondence between informants' reports of child and adolescent mental health: Clinical relevance across cultures and assessment setting*. Symposium to be conducted at the meeting of the Association for Psychological Science, Chicago, IL.

Marsh, J. K., & Zeveney, A. (2016, March). Thinking about treatments in the mental health domain. In N. S. Kim (Chair), *Beliefs and judgments about mental disorders: Causes, consequences, and treatment*. Symposium conducted at the meeting of the Eastern Psychological Association, New York, NY.

Zeveney, A., Marsh, J. K., & De Los Reyes, A., (2016, March). Lay integration of mental health information from discrepant sources. Talk presented at the meeting of the Eastern Psychological Association. New York City, NY.

Marsh, J. K., & Zeveney, A. (2015, November). The influence of category membership on causal essence beliefs. Talk presented at the 55th Meeting of the Psychonomic Society. Chicago, IL.

Marsh, J. K., & Zeveney, A. (2015, July). Naïve beliefs about intervening on causes and symptoms in the health domain. Poster presented at the 37th Annual Conference of the Cognitive Science Society. Pasadena, CA.

Marsh, J. K., & Zeveney, A. (2015, March). The influence of category membership on causal essence beliefs. Talk presented at the meeting of the Eastern Psychological Association. Philadelphia, PA.

Zeveney, A., & Marsh, J. K. (2014, November). The influence of multiple routes to category membership on category beliefs. Talk presented at the 55th Meeting of the Psychonomic Society. Long Beach, CA.

Zeveney, A. (2014, April). The impact of multiple causes on categorization. Talk presented at the 29th Lehigh Valley Undergraduate Research Conference for Psychology. Center Valley, PA.

RESEARCH FUNDING

Lehigh University Student Forum Research Grant Amount: \$50	2015
College of Arts and Sciences Undergraduate Research Grant Amount: \$250	2013

TEACHING EXPERIENCE

Teaching Assistant, Personality Psychology

- Leading weekly small group discussion section of 15 students.
- Designing assignments and in-class activities and grading assignments.
- Holding weekly office hours and meeting with students.

Teaching Assistant, Abnormal Psychology

- Held weekly office hours and met with students.
- Lead exam review sessions.
- Graded student exams.

COMPUTER SOFTWARE

Proficient in:

SPSS

Microsoft Office

Qualtrics

Knowledge of:

SAS

MPlus

REFERENCES

Dr. Jessecae Marsh, Assistant Professor, Psychology Department, Lehigh University

Email: jessecae.marsh@lehigh.edu

Phone: 610-758-2941

Dr. Lucy Napper, Assistant Professor, Psychology Department, Lehigh University

Email: lucy.napper@lehigh.edu

Phone: 610-758-4689

Dr. Bridget Dever, Assistant Professor, School Psychology Department, Lehigh University

Email: bdever@lehigh.edu

Phone: 610-758-3210