Mental health assessment: Inference, explanation, and coherence

Paul Thagard PhD, Distinguished Professor Emeritus1 | Laurette Larocque PhD, Registered Psychotherapist2

1 Department of Philosophy, University of Waterloo, Waterloo, Ontario, Canada
2 Private practice, Guelph, Ontario, Canada

Correspondence
Paul Thagard, Department of Philosophy, University of Waterloo, Waterloo, Ontario, Canada. Phone: 519-725-3036. Email: pthagard@uwaterloo.ca

Abstract
Mental health professionals such as psychiatrists and psychotherapists assess their patients by identifying disorders that explain their symptoms. This assessment requires an inference to the best explanation that compares different disorders with respect to how well they explain the available evidence. Such comparisons are captured by the theory of explanatory coherence that states 7 principles for evaluating competing hypotheses in the light of evidence. The computational model ECHO shows how explanatory coherence can be efficiently computed. We show the applicability of explanatory coherence to mental health assessment by modelling a case of psychiatric interviewing and a case of psychotherapeutic evaluation. We argue that this approach is more plausible than Bayesian inference and hermeneutic interpretation.

KEYWORDS
assessment, coherence, diagnosis, explanation, inference, psychiatry, psychotherapy

1 INTRODUCTION

Mental health professionals (MHP) include psychiatrists, clinical psychologists, psychotherapists, counsellors, psychiatric nurses, and some social workers and family physicians. When patients present with mental problems, MHP need to assess whether the patients are mentally ill by identifying disorders or diseases that would explain their symptoms. For example, depression has symptoms such as persistent sadness and hopelessness, and schizophrenia has symptoms such as hallucinations. Assessment is a complex inferential process that takes into account many factors besides symptoms, including patients’ descriptions of their family and work situations.

The primary purpose of assessment is to suggest treatment to alleviate the patients’ problems, ranging from medication such as antidepressants to different forms of psychotherapy. Choice of therapies is another complex inferential process that takes into account the assessment of the patients’ mental state and estimation of what therapies are most likely to be effective. Whereas assessment is an inference about what explains the patients’ difficulties, treatment is a decision about how to deal with those difficulties.

Understanding assessment and treatment requires both logic and psychology. Logic concerns how people ought to make inferences, while psychology concerns how people actually do make inferences. Here are some challenging questions about the logic and psychology of mental health assessment:

- What forms of inference do MHP use to identify mental problems?
- Does mental health assessment follow the same logic as medical diagnosis in general?
- What cognitive processes are used by MHP in inferences concerning assessment and treatment?

We will answer these questions by considering a range of possible ways in which assessment inferences might work, including deduction, pattern matching, and causal reasoning. We argue that deduction and pattern matching are too simple to capture the reasoning involved in assessment, which requires identification of causal networks. Assessment is a kind of inference to the best explanation, where disorders are hypotheses that explain symptoms. Inference to the best explanation is performed by explanatory coherence, a mental and logical process that can be performed by artificial neural networks.1 Similar accounts apply to medical diagnosis in general and to other causal inferences in medicine such as those found in epidemiology and in medical research. We contrast this account of medical inference with Bayesian and hermeneutic approaches.
We use a detailed case study of a psychiatric interview leading to a diagnosis of severe depression. We analyse the structure of the causal reasoning it requires, which we interpret using the theory of explanatory coherence. That theory is paired with a computational model, which we use to simulate the psychiatrist’s reasoning. We also model a psychotherapist’s thinking in assessing a client with illness anxiety disorder.

2 | METHODS

The methods used in this article are case studies, logical analysis, and cognitive modelling. A cognitive model requires specifying a mental process with sufficient precision that it can be simulated on a computer.\(^2\)\(^3\) We use the theory of explanatory coherence that is implemented in a computer program called ECHO.

Our primary case study is an annotated 60-minute interview conducted by the psychiatrist Shawn Christopher Shea.\(^4\) The fictitiously named patient is Gary Whitman, a 63-year old, white, male, married, American military veteran. In his interview, he describes many problems such as feeling neglected by his grown-up children, becoming progressively more depressed over several years, sometimes imagining suicide, and having chronic headaches that he attributes to head trauma while in the Navy. The interview reveals numerous other symptoms of depression such as decreased energy and concentration. Accordingly, Shea diagnoses him with severe depression and prescribes an antidepressant drug, also recommending that he continue psychotherapy with another practitioner. Our key question concerns how to understand the inferences made by Shea concerning diagnosis and treatment. Shea’s discussion is compatible with other accounts of mental health interviewing.\(^5\)\(^6\)

Awareness of the importance of hypotheses goes back to the Renaissance. In the nineteenth century, Charles S. Peirce gave the name “abduction” to inferences to explanatory hypotheses, covering both the invention of hypotheses and their assessment.\(^7\) In the 1960s, the acceptance of hypotheses was dubbed “inference to the best explanation”\(^8\)\(^9\). Gilbert Harman argued that such inference requires considering how many hypotheses and pieces of evidence fit together in accord with explanatory coherence.\(^10\) In the 1980s, Thagard realized that explanatory coherence can be computed by artificial neural networks that take into account numerous constraints among hypotheses and evidence, and many applications followed to scientific and legal reasoning.\(^1\)\(^2\)\(^11\)

Table 1 presents simplified principles of explanatory coherence, chosen because of their fit with the practice of scientists such as Darwin in The Origin of Species.\(^12\) In the Whitman case, the data (Principle E4) are the results of observations, especially what Whitman says in his interview, but also Shea’s observations of behaviour such as avoiding eye contact and crying. The hypotheses are conjectures about what might be causing the data, for example, that Whitman is severely depressed, which explains why he says he is depressed; alternatively, he might be lying. Principle E2 says that hypotheses cohere with what they explain, so the hypothesis that he is depressed coheres with the data that he says that he has decreased energy and concentration. Hypotheses can be stacked up in complex causal networks; for example, Whitman is depressed because his children neglect him, which could have other hypothetical causes too. In accord with Principle E1, the coherence relation is symmetrical: hypothesis and data cohere with each other. In contrast, the probability of a hypothesis given data is usually very different from the probability of data given evidence.

Principle E3 is not obviously relevant to the Whitman case, because Shea does not mention any analogous cases, but MHPs often recognize that a current client is similar in important ways to a previous one and therefore draw analogies. Principles E5 and E6 establish incoherence relations between hypotheses that are flat-out contradictory or merely competing to explain the same data. For example, the hypothesis that Whitman is depressed contradicts the alternative hypothesis that he is only faking depression, while the hypotheses that he has anxiety or grief merely compete to explain some of the same data such as his failure to make eye contact: is possible that he is depressed, anxious, and grieving.

Principles E1 to E6 establish complex networks of data, explanations, and competing hypotheses at different levels. Principle E7 directs how to determine what to believe and what not to believe, based on how well a proposition (hypothesis or piece of evidence) fits with everything else. For example, Shea’s hypothesis that Whitman suffers from clinical depression should fit with all the data and outcompete alternative hypotheses.

To determine overall coherence, the computer program ECHO uses a neural network algorithm for approximately maximizing coherence. ECHO represents each proposition by a unit, a simplified artificial neuron that is connected to other units by excitatory and inhibitory links. As in real neurons, an excitatory link is one that enables one neuron to increase the firing of another, whereas an inhibitory link decreases firing.

In the Whitman example, we can represent the hypothesis that he is depressed by a unit called DEPRESSED and the evidence that he has low energy by a unit LOW-ENERGY. Then whenever Principles E2 and E3 establish relations of coherence between 2 propositions, the units that represent the propositions get excitatory links between them.
So DEPRESSED and LOW-ENERGY have an excitatory link between them that is symmetric in accord with Principle E1. Principle E4 is implemented by making an excitatory link between a special unit EVIDENCE and any unit such as LOW-ENERGY that represents a proposition based on observation. Principles E5 and E6 that establish incoherence, when 2 hypotheses are incoherent, eg, DEPRESSED vs. LIES-DEPRESSED, then the units that represent them will get an inhibitory link between them. In principle, weights on links could vary to represent differences in the perceived strengths of explanations and conflicts, but information about these differences is rarely available, so default weights are used, positive for excitation and negative for inhibition.

The acceptability of a unit is represented by its activation, corresponding roughly to the firing rate of a real neuron. Just as firing rates of neurons are determined by their excitatory and inhibitory neurons, the activation of units in ECHO are determined by their excitation and inhibition. When the network settles (ie, activations stabilize), the resulting activations (positive or negative) indicate whether the hypotheses and data represented by the units are accepted or rejected. The test of the theory of explanatory coherence is whether examples such as the Whitman interview can be plausibly modelled by using the program ECHO.

3 | RESULTS

3.1 | ECHO simulation of the Whitman case

ECHO models Shea’s assessment of Whitman given the input in the appendix in the Supporting Information. Figure 1 shows a simplified picture of the structure of the ECHO simulation, hypotheses indicated in capitals, and data in lower cases. There are many pieces of data gathered from the interview, such as that Whitman cries, says he is depressed, has low energy, and lacks concentration. Shea considers not only the hypothesis of depression that can explain these data but also alternative hypotheses such as that he has bipolar disorder or is psychotic. These alternatives are contraindicated by Whitman’s reports that he has no episodes of bizarre energy (ie, no mania which would be part of bipolar disorder) and no hallucinations that would be part of psychoses such as schizophrenia.

Moreover, Shea’s questions were aimed also at determining why Whitman is depressed, revealing several explanations. Whitman’s own explanations of his mood include missing his son in Arizona, although further questioning reveals that he has considerable contact with other children, leading Shea to hypothesize that Whitman overgeneralizes, which could be an effect as well as a cause of depression. ECHO has no problem accommodating such feedback loops, unlike the Bayesian approach discussed below. Other factors explaining Whitman’s depression could be that he feels guilty about criticizing his devoted wife and that he has headaches that have led to conflicts with the Veteran’s Administration concerning whether injuries during his Navy service are responsible for his headaches.

In sum, Shea’s inference that depression is the best explanation of Whitman’s problems is based on coherence coming from 3 directions. First and most important, depression explains numerous symptoms such as crying and lack of concentration. Second, Whitman’s depression is in turn explained by various problems including missing his children and having headaches. Third, the hypothesis of depression successfully competes with alternative hypotheses such as bipolar disorder and psychosis that conflict with evidence. ECHO integrates these considerations and computes that depression is the correct assessment and the alternatives should be rejected, taking less than a second to settle after 134 rounds of activation adjustment based on excitatory and inhibitory links. The hypothesis of depression (H1 in the appendix) ends up with positive activation, whereas bipolar disorder, psychosis, and obsessive-compulsive disorder end up with negative activation, signifying rejection.

3.2 | Other simulations of medical inference

The Whitman simulation is compatible with other applications of ECHO. We have also simulated the assessment by Laurette Larocque of one of her cases concerning illness anxiety disorder (hypochondria), incorporating the 3-directional coherence identified in the Whitman case. This case concerns a woman with numerous symptoms including constantly checking her moles for cancer, overworking, and insomnia. The hypothesis of illness anxiety is accepted for its explanatory coherence because it explains clinical observations and because the hypothesis itself can be explained by the client’s past history of fear of death, taking into account the alternative hypothesis of depression.

ECHO has also been applied to a simple case of medical diagnosis outside the realm of mental health, concerning vertigo. Given symptoms such as dizziness and trouble walking, it infers that a patient has benign paroxysmal positional vertigo rather than Ménière’s disease or labyrinthitis because the patient displays nystagmus and is helped

---

**FIGURE 1**  Simplification of coherence network for Shea’s diagnosis of Whitman. Hypotheses are in caps, while data (evidence) are in lower case. Positive constraints (excitatory links) between hypotheses and evidence are shown by solid lines, and negative constraints (inhibitory links) are shown by dotted lines
by the Epley manoeuvre. This example suggests that medical diagnosis in general, like mental health assessment, can be understood as inference to the best explanation based on explanatory coherence.

Finally, ECHO successfully models scientific reasoning in the medical realm concerning the causes of disease. Explanatory coherence explains the initial rejection and eventual acceptance of the theory that most stomach ulcers are caused by Helicobacter pylori infection.\(^{13,14}\) ECHO has also been used to simulate the recent conclusion that the rapid increase in cases of microcephaly in newborns is the result of spread of the Zika virus.\(^ {15}\) It is thus plausible that explanatory coherence provides a unified account of medical inference encompassing mental health assessment, disease diagnosis in general, and reasoning concerning medical causation.

### 3.3 Comparison with other accounts

Just as mental health assessment requires comparison of competing explanations, explaining medical inference requires considering alternatives to explanatory coherence. We examine 4 alternatives for explaining how MHPs make assessments: deduction, pattern matching, Bayesian inference, and hermeneutic interpretation.

Medical diagnosis may sometimes be carried out deductively using rules such as “If a patient has jerky, uncoordinated body movements and a mutation in the gene Huntingtin, then the patient has Huntington’s disease.” Most diseases, however, including mental illnesses, are too complicated in their symptoms, causation, and alternative explanations to be assessed by deductive application of universal rules.

Whitman had so many of the standard indicators of depression that it might seem that all that Shea needed to do is match the symptoms against the criteria listed in the Diagnostic and Statistical Manual (DSM) of the American Psychiatric Association.\(^ {16}\) For depression, the criteria are as follows:

1. Depressed mood most of the day
2. Markedly diminished interest or pleasure
3. Significant weight loss or gain
4. Insomnia or hypersomnia
5. Psychomotor agitation or retardation
6. Fatigue or loss of energy nearly every day
7. Feelings of worthlessness or guilt
8. Diminished ability to think or concentrate
9. Recurrent thoughts of death or suicide

According to DSM 5, a diagnosis of major depressive disorder is appropriate if a patient has either a depressed mood or a loss of interest for more than 2 weeks, and at least 5 of the 9 symptoms. Whitman’s interview revealed that he had been suffering for years from 1, 2, 4, 6, 7, 8, and 9, so diagnosis of depression is immediate. So it might seem that there is no need for a complex inference using explanatory coherence.

Perhaps there are some MHP who simply match symptoms against the DSM, but sophisticated mental health assessment is more complicated. First, like differential diagnosis in medicine more generally, assessments need to distinguish disorders from others that present similar clinical features. The DSM features for depression could be indications of other disorders such as bipolar disorder, schizophrenia, and anxiety, and mental health assessment needs to pick out the best explanation, not just a good enough match. So the competitive aspect of explanatory coherence is needed.

Second, the causal structure shown in Figure 1 and the appendix is important not just for diagnosing patients but also for understanding their conditions and devising effective treatments. The most effective treatments for depression use a combination of antidepressant medications and psychotherapy.\(^ {17}\) Understanding the causal origins of someone’s depression is not very important for figuring out what medications to prescribe. But origins are crucial for planning psychotherapy to deal with patients’ perceptions and misperceptions of their life situations. For example, a psychotherapist aware of Whitman’s problems concerning his children, his wife, and the Veteran’s Administration could set out to discuss these issues with him to change his cognitions and emotions in ways that alleviate his depression and resulting symptoms. The purpose of mental health assessments is not just to come up with true or probable hypotheses but to find useful ones that guide treatments that improve the patients’ lives. This pragmatic aspect of mental health assessment is discussed below.

In principle, the evaluation of competing hypotheses concerning mental disorders could be carried out by using Bayes’ theorem, which says that the probability of a hypothesis given evidence is equal to the prior probability of the hypothesis times the probability of the evidence given the hypothesis, all divided by the probability of the evidence. In symbols, this is \(P(H|E) = P(H) * P(E|H)/P(E)\). Then the probability that Whitman is depressed given all the evidence accumulated in the interview would be calculated by multiplying the prior probability of depression times the probability of the evidence given depression, divided by the probability of the evidence.

There are several problems with applying Bayesian calculations to human inference.\(^ {2-18,19}\) First, the relevant probabilities are often not available. What is the prior probability that Whitman is depressed? Equally mysterious is the probability of all of his symptoms given depression, eg. \(P(lack-of-energy|depression)\). In some rare areas of medicine, there may be large data bases that provide probabilities linking diseases and symptoms, but not in mental health.

Second, the interpretation of probability is problematic. The 2 most common interpretations are that probabilities are frequencies or subjective degrees of belief. Given the lack of frequency information, subjective degrees of belief are the obvious choice for mental health, but numerous studies show that people do not naturally think in terms of probability.\(^ {20}\) Moreover, we should want some objectivity in the judgement that the hypothesis of depression is more probable than the alternatives, but it is not clear how subjective degrees of belief can yield such objectivity.

Third, there are numerous technical problems involved in the computation of probabilities using Bayesian networks of the sort needed to capture the causal connections in the ECHO simulations.\(^ {21}\) For example, the feedback loop from depression to overgeneralization to depression cannot be computed in Bayesian causal networks that prohibit cycles. Contention that MHP do or should use Bayesian reasoning...
should be tested by modelling of realistic cases using programs such as JavaBayes. For these 3 reasons, the Bayesian approach to mental health assessment and medical diagnosis is inferior to the explanatory coherence approach.

As a final alternative to explanatory coherence as an account of mental health assessment, consider the hermeneutic approach that claims that understanding a patient's condition is analogous to interpreting a text by constructing a narrative.\textsuperscript{22-24} Narratives can be constructed by listening, empathy, detective work, and making meaning.\textsuperscript{25} The major problem with this approach is that not all narratives are equally good. For example, I might interpret the text of Moby Dick by conjecturing that the large whale was an alien from another solar system, but this interpretation fits poorly with the rest of the story. Similarly, understanding in medicine should be more than just a subjective feeling of grasping someone's condition, but instead should be based on a good causal explanation tied to the available evidence.

Hence, hermeneutics in medicine, if it aims at understanding and helping patients, needs to look for the best available interpretations of the patient's conditions. Achieving a good narrative requires inference to the best explanation. The theory of explanatory coherence stated in the principles in Table 1 is the most thorough account currently available of inference to the best explanation and naturally extends to cover textual interpretation and narrative evaluation based on listening to accumulate data about the patient, detective work to develop hypotheses that explain the data, and making meaning by causal explanation.

One of this paper's reviewers suggested that a hermeneutic approach might be especially sensitive to culture-related diagnostic issues. But explanatory coherence can be culturally sensitive by including evidence about a patient's cultural background and considering alternative causal hypotheses that link disorders with symptoms in different ways depending on social expectations. For example, emotions are expressed differently in different cultures, so the emotional disturbances found in all mental illnesses may present differently and therefore require different explanations.

What about empathy? Empathy has several modes, including using mirror neurons to enter into a mental state similar to another person and using analogy to put yourself in someone else's shoes.\textsuperscript{26,27} An MHP can empathize with a patient by noticing systematic similarities between the patient's condition and some previous episode in the professional's experience. The theory of explanatory coherence can incorporate such empathic analogies through Principle 3, Analogy. For example, a professional might think: "If I had the kind of family and work problems that Whitman has, then I would also be depressed and lacking in energy." However, this empathy by analogy cannot serve as a stand-alone inference but should be part of a complex inference that takes into account alternative hypotheses, such as that Whitman is suffering from grief. Therefore, hermeneutics of patient care, done well, results from explanatory coherence.

4 | DISCUSSION

One of the reviewers of this paper pointed to a different way of thinking about diagnosis as a sorting of symptoms rather than an explanatory inference. Perhaps “depression” is just a category summarizing a set of symptoms, rather than a causal explanation of why people have those symptoms. This view of diagnosis is consistent with the symptom-counting practice of DSM 5, but is inferior to the explanatory inference view in several respects. First, mere pattern matching encourages just counting symptoms rather than assessing, which ones are most revealing of underlying causal conditions. Second, disorders often overlap, and a good explanation of symptoms sometimes requires combining disease diagnoses. Third, disorders construed as categories lack causal underpinnings, but psychiatry is now moving toward mechanism-based understanding of mental illness through the Research Domain Criteria project of the US National Institute of Mental Health. Fourth, symptom-based categorization provides limited clues to how to treat people using deep therapies that deal with multiple layers of causality of disease. Deciding what treatments to provide for a mental disorder is not an automatic intervention, but requires inference concerning the best of various alternatives for helping people with their illnesses.

Medical thinking is not just theoretical inference about what is true, but is also directly relevant to practical inference about actions to help people with their health problems. Theoretical and practical inference would be independent if MHP proceeded by first using Bayesian reasoning to select hypotheses about disorders, and then using economics-style maximization of expected utility to decide what to do about them. This procedure is unavailable because of lack of knowledge of the relevant probabilities and utilities, and also runs counter to the goal-directed practice of health professionals.

There are several ways in which mental health assessment is affected by the pragmatics of health care. First, accumulation of data during the initial interview and subsequent sessions is appropriately driven by practical concerns about treatment. Mental health professionals are concerned from the start to find effective treatments, so the course of questioning is affected by concerns about how a patient might be helped. For example, when Shea interviewed Whitman, depression was immediately an issue so some of Shea's questions were directed at determining likely treatments such as medication and psychotherapy. The assessment interview also has to deal with practical issues such as arranging appointments and payments.

Second, both accumulation of evidence and mental health assessment are influenced by concerns about dangerous outcomes such as suicide. When patients like Whitman describe suicidal thoughts, MHP appropriately acquire information concerning whether suicide is likely, which can require immediate measures such as hospitalization.

Third, the purpose of the initial interview and subsequent sessions with MHP is not just theoretical and practical inferences about assessment and treatment. Research has shown that a major factor in treatment success for mental disorders is the formation of a therapeutic alliance based on engagement, respect, rapport, and trust between MHP and their patients.\textsuperscript{26} Hence, MHP need to ensure that their investigations and recommendations have the practical result of building a good relationship with clients who will then be more likely to engage actively in the treatment. Even one session can help move mental states from fear to hope.

Fourth, mental health assessment is not a 1-time inference occurring in the first interview, but requires ongoing revision based on additional information obtained by patients' statements and by
the perceived effectiveness of treatments. For example, if Whitman's depression lifts after a few weeks on antidepressant medication, the recovery confirms the initial diagnosis. Conversely, treatment failure suggests the need for another diagnosis, for example, when a physician treats an infection with antibacterial medication that fails to help. Over the course of treatment, MHP and physicians may dynamically change their diagnoses as evidence, life events, and treatments move forward.

Epistemologists debate whether "pragmatic encroachment" of practical concerns on inference is legitimate. The 4 issues of focused gathering of data, consideration of dangerous outcomes, therapeutic alliance, and inferences drawn from treatment success or failure show that reasoning by MHP is appropriately goal-driven, in ways better described as pragmatic enhancement rather than encroachment.

The theory of explanatory coherence says nothing about these pragmatic issues, but it is part of a broader theory of coherence that includes practical inferences about what to do, including emotion-driven inference. More investigation is needed to understand pragmatic enhancement in medical inference from a coherentist perspective.

In conclusion, mental health assessment is a medically important task whose logic and psychology have received little attention. We have argued that the theory of explanatory coherence is better than alternative accounts that include deduction, pattern matching, Bayesian inference, and hermeneutic interpretation. Thanks to its implementation in the computer program ECHO, explanatory coherence successfully models realistic cases of reasoning concerning disorders that explain patients' mental symptoms. Future work is required to integrate explanatory coherence with practical inferences and pragmatic enhancement.

ACKNOWLEDGEMENTS
For helpful suggestions, we are grateful to Steve Bank, Ramesh Prasad, Miriam Solomon, and 2 anonymous referees.

ORCID
Paul Thagard
http://orcid.org/0000-0002-6594-4761

REFERENCES

SUPPORTING INFORMATION
Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Thagard P, Larocque L. Mental health assessment: Inference, explanation, and coherence. J Eval Clin Pract. 2018;1–6. https://doi.org/10.1111/jep.12885