Explaining Mental Illness
Paul Thagard
University of Waterloo

Outline
1. Mental illness
2. What is explanation?
3. What is mind?
4. What is disease?
5. Breakdowns in mental mechanisms
6. Depression

3. Analysis of Mental Illness

1. Exemplars: depression, anxiety, schizophrenia, bipolar disorder, autism, etc.
2. Typical features: problems in thinking, emotion, social functioning, etc.
3. Explains: cognitive, emotional, and social dysfunctions
   Explained by ???


Styles of Explanation: Different Patterns

1. Explanation is telling a story that answers a question. Narrative, e.g. Freud
2. Explanation is explaining away, elimination. E.g. Szasz, Laing, Foucault
3. Explanation is deduction from scientific laws.
4. Explanation is showing how something results from a causal mechanism. ✔
Mechanism

Mechanism = system of interconnected parts whose interactions produce regular changes. Salmon, Bechtel, Craver, Darden

Changes may be emergent, i.e. belonging to wholes but not to parts because they result from interactions of parts. Wimsatt, Bunge

Mechanisms break: parts, interactions.

Produce=cause.

3-Analysis of Causality

1. Exemplars: pushes, pulls, motions
2. Typical features:
   a) Sensory-motor-sensory patterns – infants
   b) Regularities: children
   c) Manipulations: children
   d) Statistical dependencies + causal networks + connections: scientists, philosophers
3. Explains: why things happen, why interventions work. Explained by ???

Theories of Mind

1. Mind=soul (dualism, e.g. Descartes)
2. Mind=nothing (behaviorism, e.g. B. F. Skinner)
3. Mind=computer (functionalism, e.g. Turing)
4. Mind=brain (identity theory, e.g. J.J.C. Smart)

Current: Mental processes are brain mechanisms.✔

Semantic Pointers (Eliasmith 2013)

Semantic pointers are patterns of neural firing that:
1. provide shallow meaning through symbol-like relations to the world and other representations;
2. expand to provide deeper meaning with relations to perceptual, motor, and emotional information;
3. support complex syntactic operations;
4. help to control the flow of information through a cognitive system to accomplish its goals.
Emotions are Semantic Pointers

Emotion = bind (concept or belief, cognitive appraisal, physiological perception)

Example: being happy to be in Paris = bind (Paris, appraisal, physiology)

Emotions are brain mechanisms.
Thagard and Schröder, Emotions as Semantic Pointers, 2014.

What is disease, illness, disorder?

1. Disease=set of symptoms, e.g. DSM
2. Disease= social construction: values
3. Disease=physiological malfunction-Boorse
4. Health= functional efficiency with respect to fitness (Hausman, 2012)
5. Disease=breakdown in mechanisms (Thagard, Pathways to Medical Discovery, 2003; What is a Medical Theory, 2006; Mental Illness from the Perspective of Theoretical Neuroscience, 2008)

3-Analysis of Disease

1. Exemplars: influenza, cancer, scurvy, arthritis, heart failure, Alzheimer’s ...
2. Typical features: symptoms, malfunctions, harms
3. Explains: why people have harmful symptoms
Explained by: breakdown in physiological mechanisms
Causes of Mental Illness

1. **Neural**: breakdowns in neurons (e.g. Huntington’s) and neural interactions (e.g. Parkinson’s)

2. **Molecular**: breakdown in molecular pathways, e.g. dopamine, serotonin

3. **Psychological**: breakdown in representations, inferences, and emotions, e.g. Capgras syndrome

4. **Social**: stresses such as childhood abuse, unemployment, relationship problems

---

Thagard (2014), The self as a system of multilevel interacting mechanisms, *Philosophical Psychology.*
### Molecular Mechanisms

- Neurotransmitters, hormones, epigenetics
- Evidence based on effectiveness of antidepressant medications:
  - Serotonin reuptake inhibitors (Prozac, Zoloft)
  - Serotonin + norepinephrine (Effexor)
  - Serotonin + dopamine (Wellbutrin)
  - Monoamine oxidase inhibitors

### Neural Mechanisms

- Prolonged sadness = prolonged binding of negative appraisal and physiology to all situations.
- Why do antidepressants take weeks to work?
  - Neurogenesis (brain-derived neurotrophic factor)
- Explanation pattern: stress -> increased cortisol -> decreased neurogenesis in hippocampus -> decreased mental flexibility -> depression
- Why do deep brain stimulation and electroconvulsive therapy sometimes work?

### Psychological Mechanisms

- Why does cognitive therapy help?
  - Cognitive therapy adjusts beliefs, goals, and emotions.
  - Helps with re-appraisals concerning loss, rejection, self-worth, etc.
  - Avoid rumination (repeated negative thoughts)

### Social Mechanisms

- Social causes of depression:
  1. Childhood abuse
  2. Bereavement
  3. Unemployment
  4. Relationship failures
  5. Social rumination
Implications
Depression results from multilevel emergence of prolonged negative emotions because of breakdowns in neural, molecular, psychological, and social mechanisms.
Treatment: try to restore functioning of ALL mechanisms.

Conclusions
1. Biological explanations are descriptions of causal mechanisms.
2. Mental mechanisms are neural, but also molecular, psychological, and social.
3. Mental illnesses result from breakdowns in multilevel mechanisms.

Binding in the Brain
Synchrony: neurons fire in temporal coordination
Syntax: e.g. Shastri, Hummel
Consciousness: e.g. Crick, Engel, Scherer
Convolution: activity of neural populations becomes "twisted together": convolve.
Representations are braided together.
Eliasmith has shown how neural populations can perform convolution.
Three Mechanisms

<table>
<thead>
<tr>
<th>Parts</th>
<th>Interactions</th>
<th>Emergent result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurons</td>
<td>Excitation, inhibition, synaptic connections</td>
<td>Representation by firing patterns</td>
</tr>
<tr>
<td>Neural groups</td>
<td>Recursive binding</td>
<td>Semantic pointers</td>
</tr>
<tr>
<td>Semantic pointers</td>
<td>Interactive competition</td>
<td>Conscious experience</td>
</tr>
</tbody>
</table>

Emergence

Emergent properties are possessed by the whole, not by the parts, and are not simple aggregates of the properties of the parts because they result from interactions of parts.