

Malfunction and harm: Why the distinction  
doesn't work to ground psychiatry (B)

Kelly Alexandra Roe

2007

*Presented to the Australian Association of Philosophy conference. Hosted by the University of Auckland*

## **Situating the Topic**

- My greater project is to see how psychiatric classification can progress scientifically given their stated aims
- – Tool for researchers
- – Tool for clinicians
- The first issue that I'm concerned with is:
  - – what makes it the case that a person has a psychiatric disorder?
- This will lead in to the greater project of figuring out what the basic units of a psychiatric taxonomy should be
  - – Issue of natural kinds
  - – Issue of different kinds of causal mechanisms (e.g., genetic, neurological, cognitive psychological, social)
- Today I'm going to focus on a problem that is located within the first issue

## **Preliminaries**

- I'm not going to say anything today about what differentiates psychiatric disorder from non-psychiatric medical disorders, however
- So while psychiatry is my main area of interest I'm going to be considering bio-medical disorders more generally for the purposes of this talk
- This means that we can make use of examples from both psychiatry and general medicine
- 'Disorder' is a stand-in for related notions like 'disease', 'malady', 'sickness', 'pathology', 'illness', and so on

## Plan

- Introduce the problem
- The anti-psychiatry critique of psychiatry
- The attempt to ground psychiatry
- The two-stage view
- Wakefield's version of the two-stage view
- Problems with the malfunction and harm distinction
  - – There seem to be objective facts about whether or not an individual is harmed relative to society
  - – Dysfunctions can be behavioural as well as internal
  - – Fixing the relevant notion of function seems to involve our adopting a standard
  - – That standard seems partly determined by our values

## The Problem

- We have the intuition that certain conditions are disorders
  - – Broken legs, HIV, cancer
  - – Depression, mania, psychosis
- The issue is figuring out what justifies our regarding these conditions as disorders
- One answer might be that the conditions are included in a classification scheme
- The problem is that previous classification schemes included conditions like homosexuality and sluggish schizophrenia
- We want to know what underlying principle justifies our including certain conditions in a classification scheme and what justifies our excluding others

## The Critique as a Motivation for the Grounding Project

- Anti-psychiatrists maintain that there is no more to mental disorder than social and / or moral norm violation
- In the face of the anti-psychiatry critique there has been an attempt to ground psychiatry in medicine
- The two-stage view is the most popular view of how the grounding should proceed
- According to the two-stage view there are two individually necessary and jointly sufficient conditions for disorder
  - – Malfunction / Dysfunction (objective – to be discovered by science)
  - – Harm( normative – varies across cultures)
- According to the two-stage view one can separate out matters of value from matters of fact and ground psychiatry solely by way of matters of fact

## Grounding Psychiatry: The ‘Naturalization Cascade’

- Mental Disorder is a certain kind of Physical Disorder
- - Psychiatry (‘Mental Disorder’, ‘Mental Illness’)
- - Disorder = Dysfunction + Harm\* (Medicine (‘Disorder’, ‘Illness’, ‘Disability’ Malady’ etc))
- Functions and Malfunctions = Physical Properties and Processes
- - Biology (‘Function’, ‘Malfunction’)
- - Physics (Physical Properties and Processes)
- More needs to be said about the malfunction and harm distinction
- I’m going to focus on Wakefield’s particular version of the two-stage view because he is clearer than most and because he has been the most vocal advocate

- I'll offer a reconstruction of his main argument for his particular version as a way into his view. This will help me illustrate some of the problems

## Wakefield's Version of the Two-Stage View

- P1) It is a conceptual truth of the bio-medical notion of disorder that disorder is a result of an internal dysfunction that results in harm to persons (where dysfunction is to be understood in some pre-theoretical sense)
- P2) It is a conceptual truth that there is an empirical process that fixes the functions and hence dysfunctions
- P3) Scientists have discovered that the relevant process for fixing functions and dysfunctions is evolution by natural selection
- C) Disorders are thus failures of an internal mechanism to perform its evolutionary function (that results in harm to persons)

## Malfunction and Harm

- The idea is that malfunction is internal to the person and it is objective (to be discovered by science)
- And that harm is a feature of behaviour and / or the effects of behaviour and is normative (to be determined by our social and / or moral values)
- And that intuitively these can come apart:
  - Harm without malfunction
    - – Never being taught how to read vs reading disorder
  - Malfunction without harm
    - – Gourmand lesion

## Malfunction and Harm

- The notion of harm is meant to be a stand-in for the normative aspect of disorder

- Not much has been said about it other than that it is to persons, that it is a feature of behaviour, that it is determined by our social and / or moral values, and that it is normative (not objective)
- Still, those do seem to be a number of substantive claims
- The focus has been on characterizing dysfunction as being to internal parts of
- This is because the grounding project is supposed to proceed by way of dysfunction persons, and objective (non-normative)

## Harm

- One might have a number of concerns with the way that harm has been characterized. In particular we might wonder:
  - – Is there an objective aspect to whether an individual and / or society is harmed?
  - There can be objective facts about what certain societies do and do not value.
  - This can be the subject matter of scientific investigation (e.g., sociology, psychology).
  - There could thus be facts about whether an individuals behaviour is harmful in relation to the values of society.
  - So, for example, there could be a fact that a person with Gourmand lesion is helped rather than harmed (relative to a society that values gourmets) and that a person who can't read is harmed (relative to a society that values reading)

## Dysfunction

- I now want to turn to problems with the dysfunction criterion as the majority of the debate has focused on this aspect
- Wakefield maintains that dysfunctions need to be internal to the person
- That scientists have discovered that evolution by natural selection is the

- That there are objective facts about functions and malfunctions that are determined by science quite independently of our values
- I want to dispute these three claims relevant process for fixing functions and dysfunctions

## Why Do Functions and Dysfunctions Have to be Internal?

- Functions and dysfunctions have been predicated of various things:
  - – Processes, states, mechanisms
  - – Effects of processes, states, mechanisms
  - – Behaviours or traits
- Inner Inner Outer
- Process State
- Mechanism
- Output
- Effect Trait
- Behaviour

## What Fixes Functions and Dysfunctions?

- There are at least four broadly different accounts of function fixing:
  - – Aristotelian Teleological (Megone) – Statistical (Boorse)
  - – Evolutionary (Wakefield)
  - – Systemic (Murphy)

## Aristotelian Teleological (*f*TELOS)

- Forward looking, teleological, and purposive
- For example, watches are for keeping time and if a watch doesn't keep time then it is dysfunctioning

- What seems relevant is that the watch was designed by an agent with a certain intention
- Aristotle thought the function of a person was reason
- A common view of mental disorder is that people with mental disorders are irrational e.g., delusions

## Statistical (*f*STAT)

- Boorse maintains that we begin by identifying the relevant reference class by way of species / gender / age
- We then assign functions and dysfunctions on the basis of statistical mean
- Dysfunctions can thus be measured in standard deviations from the mean

## Evolutionary (*f*EVO)

- The evolutionary notion of functions fixes functions by how much a trait contributes towards evolutionary fitness / expected reproduction
- E.g., ‘the functions are whatever effects of past tokens resulted in their surviving and reproducing such that there are presently existing tokens’
- Or, on the propensity view ‘the functions are whatever effects of present tokens will result in their surviving and reproducing into future generations’

## Systemic (*f*SYST)

- According to the systemic notion of function we need to begin by specifying some relevant output of a system
- – E.g., we want to explain how the circulatory system circulates blood / nutrients
- The functions of the components of the system are then fixed in virtue of the role they play in producing the relevant output



- – E.g., ‘the function of the heart with respect to the circulation of nutrients is to function as a pump’
- There has been much controversy over whether these are simply different notions of function or whether one notion can be explicated such that some or all of the others can be derived from it.
- This is especially the case with systemic and evolutionary functions as people have attempted to provide a unified account of function in biology.
- This is also the case with Aristotelian teleological and evolutionary functions as people have attempted to naturalize intentionality and rationality
- At first glance these notions of function seem to be different, however (they would differ in their assignment of functions in at least some cases)
- – E.g., Millikan on how evolutionary functions can come apart from statistical functions
- If the different notions of function deliver different verdicts as to what the functions and dysfunctions are then a defender of the dysfunction criterion would need to commit to a particular view on what functions are relevant for psychiatry and / or medicine

## Common Features: Functions as Relations Between the World and a Standard of Evaluation

- Now it seems that all the different accounts of function seem to share a common structure
- They all provide some standard such that one can assign functions and dysfunctions to physical properties and causal processes relative to the standard
- Standards include:
  - $f_{TELOS}$  – agents intentions and / or norms of rationality
  - $f_{STAT}$  – statistical mean
  - $f_{EVO}$  – expected reproduction

- $f$ SYST – some relevant output of whole
- If we have a seemingly normative claim such as ‘that heart should be pumping blood’ then we can translate that into a description of purely physical properties and processes
- –  $f$ EVO ‘that heart isn’t doing what past hearts did that enabled them to survive and reproduce’
- –  $f$ SYST ‘that heart isn’t doing what other hearts do when they contribute to the circulatory system’s circulation of nutrients
- These claims are purely descriptive
- But there is no entailment from a completed description of physical properties and processes to what the heart should be doing – in the absence of some standard of evaluation
- Why should hearts do what past hearts did? If you value survival and reproduction then the evolutionary standard fixes the function as doing what past hearts did...
- If you value death, however, then the inverse evolutionary standard fixes the function of evolution as extinction and a functioning heart would be one that made death more probable...

## Is it Really Normative?

- There can of course be facts about what norms are or are not endorsed by a given society
- There can of course also be facts about what norms a given society should or should not adopt – relative to some interest or other
- There might similarly be facts about what standard is relevant for psychiatry – given its aims
- If we ask what psychiatry’s aims should be (or what norms it should adopt in a way that divorces the ‘should’ from a standard) then that seems normative, however
- It is far from obvious that any of the previous standards are the appropriate standard for fixing the subject matter of psychiatry given its aims
- The aims (once again) were to

- – Provide a list of conditions that is useful to researchers
- – Provide a list of conditions that is useful to clinicians
- While there might be facts about what conditions are readily identifiable and treatable the question of whether we ‘should’ treat it or whether an individual should be helped seems (at least partially) determined by our values

## Conclusion

- Basically what I’ve tried to do is carve a middle way between those who maintain
- – There is no more to mental disorder than certain kinds of social and / or moral norm violation
- And those who maintain
- – Science will determine who is and who isn’t disordered
- While some theorists think that science will discover what conditions really are disorders or not I think that can’t do all the work
- We equally need to get clearer on the relevant standard for psychiatry and for medicine
- Physical properties and causal processes in the absence of a standard is insufficient to fix what conditions are and are not disordered