

論文の内容の要旨

論文題目 : **Delusions as Malfunctioning Beliefs A Biological Defense of Doxasticism about Delusion**

(信念の機能不全としての妄想 生物学的な観点から妄想の信念主義を擁護する)

氏 名 : 宮園健吾

Doxasticism about delusion (DD) is the claim that delusions are beliefs. DD is widely accepted in psychiatry. In Psychiatry, the expression of DD can be found in wide-range of places from introductory textbooks, research papers to diagnostic manuals. DD has also been defended by some philosophers including Bayne & Pacherie (2005) and Bortolotti (2010). Recently, however, a number of psychiatrists and philosophers pointed out that many delusions do not behave like beliefs. In other words, many delusions do not play the causal roles that are characteristic to beliefs (let us call them “belief-like causal roles”). Egan, for instance, wrote; “delusions, it turns out, display a lot of behavior that doesn’t look terribly belief-like. [...] The role that delusions play in their subjects’ cognitive economies differs pretty dramatically from the role that we’d expect beliefs to play.” (Egan 2009, 265-266) Let us call this claim, the claim that many delusions fail to play belief-like causal roles, “causal difference thesis” or “CDT” for short.

Now, there is a clear tension between DD and CDT. If CDT is true and many delusions fail to play belief-like causal roles, how can it be the case that those delusions are beliefs? Again, if DD is true and delusions are beliefs, then how can it be the case that many of them fail to play belief-like causal roles? This tension is especially acute given the fact that (broadly) functionalist conception of belief (including functionalism (Armstrong 1968; Lewis 1980), dispositionalism (Schwitzgebel 2001; Ryle 1949), representationalism (Fodor 1987; Nichols & Stich 2003), and, presumably, interpretationism (Davidson 1984; Dennett 1989)) is influential in the philosophical literature on belief. According to functionalist conception, playing a belief-like causal role is necessary for a mental state to be a belief. Thus, there is no such thing as the belief without belief-like causal role.

One might think, on the basis of the functionalist conception of belief, that DD and CDT are incompatible with each other and, hence, at least one of them should be rejected. This “incompatibilism” is dominant in the recent literature on delusion. There are two types of

incompatibilists. First, there are those who reject DD and accept CDT (Berrios 1991; Currie 2000; Currie & Jureidini 2001, 2003; Currie & Ravenscroft 2002; Egan 2008; Frankish 2009, 2012; Hohwy & Rajan 2012; Schwitzgebel 2012; Tumulty 2011, 2012). They are “anti-DD incompatibilists”. Anti-DD incompatibilists argue that at least many delusions are not beliefs, because they fail to play belief-like causal roles. Those delusions are instead regarded as meaningless speech acts, imaginations, some in-between states, acceptance, perceptual inferences, and so on. Second, there are those who reject CDT and accept DD. They are “pro-DD incompatibilists”. Pro-DD incompatibilists try to defend the idea that delusions are beliefs by rejecting the claim that delusions fail to play belief-like causal roles (Bayne & Pacherie 2005; Bortolotti 2010, 2011, 2012; Bortolotti & Broome 2012; Reimer 2010).

The main aim of this dissertation is to develop an option that hasn't explored seriously so far, namely, a compatibilist option. Compatibilists are those who think that DD and CDT are compatible with each other. Certainly, DD and CDT are incompatible with each other according to the functionalist theories of belief. But compatibilists argue that they are compatible with each other under some different theories of belief and those alternatives theories are, at least, as plausible as functionalist ones for independent reasons.

In Chapter 1, I introduce DD and CDT with more details and present some reasons for these claims. DD is supported by the fact that (1) delusional subjects sincerely assent to the content of their delusions, (2) the truth of DD is the part of the reason why delusion is a pathological condition, (3) DD is useful in distinguishing delusion from other pathological conditions, (4) most psychiatrists, the experts on the phenomenon, regard delusions as beliefs and (5) delusional subjects, who have authoritative first personal access to their own mental states, regard their delusions as beliefs. CDT, on the other hand, is supported by numerous empirical and clinical observations suggesting that delusions are (A) extremely insensitive to evidence, (B) can be very incoherent with other beliefs, (C) can fail to evoke appropriate affective responses, and (D) can fail to cause appropriate action.

In Chapter 2, I motivate the compatibilist option by showing (a) that there is no good argument against the possibility of mental states without their distinctive causal roles, (b) that there are some examples, other than delusions, of mental states without their distinctive causal roles, and (c) that incompatibilist options (anti-DD and pro-DD) face some difficulties. On (a), I argue for instance that multiple-realizability argument, which seems to support functionalist theories of belief, doesn't rule out the possibility of mental states without their distinctive causal roles. On (b), I suggest that pain asymbolia involves the pains without pain-like causal roles, addictive desires are the desires

without desire-like causal roles, and so on. On (c), I show that anti-DD incompatibilism has at least *prima facie* difficulty in accounting for the considerations that seem to favor DD and that Bortolotti's proposal (2010), which is the most well-developed pro-DD incompatibilist proposal, is problematic for a number of reasons.

In Chapter 3, I explore the best compatibilist theory of belief. Some candidates are examined and rejected. Type-identity theory is rejected because of its well-known problem about multiple-realizability of mental states. Humean phenomenalism is rejected because of its limited scope as a theory of belief. Lewis's statistical functionalism (1980) is rejected because of its absurd consequences involving large-scale statistical changes. Bayne's normativism (2010) is rejected because of the indeterminacy as to whether a given mental state is a belief or not. The best candidate for the compatibilist theory of belief is what I call "teleo-attitude functionalism". The core idea of this theory is that beliefs are characterized in the same way that biological organs such as hearts or kidneys are characterized. For the same reason that there can be hearts that fail to pump blood, there can be beliefs that fail to play belief-like causal roles, according to teleo-attitude functionalism. This means that DD and CDT are compatible with each other according to this theory. Hearts that fail to pump blood are "malfunctioning hearts". Analogously, the beliefs that fail to play belief-like causal roles are "malfunctioning beliefs". Teleo-attitude functionalism has some important predecessors in philosophy and psychology including; Descartes's discussion on passion (Descartes 1649/1985), contemporary psychological accounts of emotion (Nesse 1990; Oatley & Johnson-Laird 1987), teleological functionalism (Lycan 1982, 1995; Sober 1985), and teleosemantics (Dretske 1986, 1991; Millikan 1984, 1989b).

Chapter 4 is the discussion of some theoretical issues about teleo-attitude functionalism. First, I discuss the issues on so-called "etiological analysis function" on which teleo-attitude functionalism rely. In particular, I argue that etiological analysis is the best option for teleo-attitude functionalists because it is a crucial part of the theory that beliefs can malfunction and etiological analysis is the most plausible analysis of function that allows for malfunction (i.e. having the function of doing F without actually doing F). Many alternative analyses of function, such as systemic analysis (Cummins 1975), have difficulty in allowing for the possibility of malfunctioning. Counterfactual analysis (Nanay 2010) does allow for malfunction but it turns out to be less plausible than etiological analysis. Second, I respond to some expected objections to teleo-attitude functionalism, including philosophical objections with imaginary cases and empirical objections involving empirical considerations.

Following Papineau (2001), I reject philosophical objections, including Swampman objection and mad-belief objection, on the ground that teleo-attitude functionalism is presented as the account of the way the world actually is. I critically examine the empirical, anti-adaptationist objections by Fodor (2001) and Gould (1991) and show that their attacks of adaptationist evolutionary psychology are based upon some theoretical confusions.

Chapter 5 explores the implications. First, I argue that DD is likely to be true, not merely compatible with CDT, according to teleo-attitude functionalism. According to the theory, a mental state is a belief just in case it has the right kind of etiological function. Since the etiological function of mental states are somehow derived out of etiological function of the mechanisms that produce and use the states (Millikan 1984), this amounts to the view that a mental state is a belief just in case it is produced by the right kind of mechanisms and used by the right kind of mechanisms. And I describe the ways in which current and future empirical studies of delusion formation can help us to determine if delusions are produced by the right kind of mechanisms and used by the right kind of mechanisms. Second, I show that my proposal provides a nice account of the fact that delusion is, unlike self-deception or wishful thinking, regarded as a pathological condition. Delusions are, in my view, malfunctioning beliefs. I argue, on the basis of this, that delusion is a pathological condition because it is harmful (i.e. negatively affect well-being and social functioning) and it involves etiological malfunctions (of some underlying mechanisms). This account is just another application of Wakefield's influential account of disorder according to which a condition is a disorder just in case it involves harmful etiological malfunction (Wakefield 1992a, 1992b).

Appendix, which is somewhat independent from the earlier chapters, discusses the process of delusion formation in detail. In particular, it examines the relationship between two prominent theories of delusion formation, namely, two-factor theory and prediction-error theory. We will argue that, on the contrary to a popular view, those two views are theoretically compatible with each other, and suggest some particular ways in which two views are actually combined together.